

# RS  
2

JC979 U.S. PTO  
10/020077  
10/30/01

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In Re the Application of : Ryo TAKAJITSUKO, et al.  
Filed: : Concurrently herewith  
For: : COMMUNICATIONS APPARATUS AND.....  
Serial No. : Concurrently herewith

Assistant Commissioner for Patents  
Washington, D.C. 20231

October 30, 2001

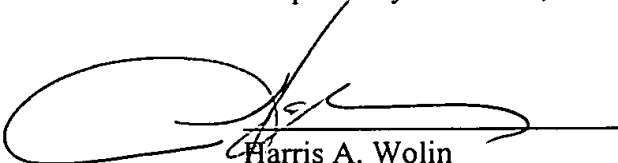
SUBMISSION OF PRIORITY DOCUMENT

S I R:

Attached herewith is **JAPANESE** patent application nos. **2001-196778** filed **June 28, 2001** whose priority has been claimed in the present application.

Any fee, due as a result of this paper, not covered by an enclosed check, may be charged to Deposit Acct. No. 50-1290.

Respectfully submitted,

  
Harris A. Wolin  
Reg. No. 39,432

ROSENMAN & COLIN, LLP  
575 MADISON AVENUE  
IP Department  
NEW YORK, NEW YORK 10022-2584  
DOCKET NO.: FUJI 19.117  
TELEPHONE: (212) 940-8800

日 本 国 特 許 庁  
JAPAN PATENT OFFICE

JC979 U.S. PTO  
10/020077  
10/30/01

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出 願 年 月 日

Date of Application:

2001年 6月28日

出 願 番 号

Application Number:

特願2001-196778

出 願 人

Applicant(s):

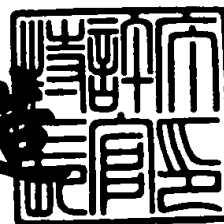
富士通株式会社

CERTIFIED COPY OF  
PRIORITY DOCUMENT

2001年 8月31日

特許庁長官  
Commissioner,  
Japan Patent Office

及川耕造



出証番号 出証特2001-3081002

【書類名】 特許願

【整理番号】 0100087

【提出日】 平成13年 6月28日

【あて先】 特許庁長官 及川 耕造 殿

【国際特許分類】 H04L 12/00  
H04Q 3/00

【発明の名称】 通信装置及び通信制御方法

【請求項の数】 5

【発明者】

【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

【氏名】 高實子 亮

【発明者】

【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

【氏名】 岡部 健一

【発明者】

【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

【氏名】 瓜生 士郎

【発明者】

【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

【氏名】 河▲崎▼ 裕哉

【特許出願人】

【識別番号】 000005223

【氏名又は名称】 富士通株式会社

【代理人】

【識別番号】 100070150

【住所又は居所】 東京都渋谷区恵比寿4丁目20番3号 恵比寿ガーデン  
プレイスタワー32階

【弁理士】

【氏名又は名称】 伊東 忠彦

【電話番号】 03-5424-2511

【手数料の表示】

【予納台帳番号】 002989

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9704678

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 通信装置及び通信制御方法

【特許請求の範囲】

【請求項 1】 異なるインタフェース間のスイッチング機能を有する通信装置において、

固定長のデータをスイッチングするメインスイッチ部と、該メインスイッチ部の入力及び出力にそれぞれ設けられた第 1 及び第 2 のバッファを具備するインタフェース部とを設けたスイッチ部を有することを特徴とする通信装置。

【請求項 2】 前記通信装置は更に、前記スイッチ部に接続され、所定のプロトコルに従った処理を行う処理部を有し、

該処理部は前記第 1 及び第 2 のバッファに接続する第 3 及び第 4 のバッファを有し、

前記第 1 のバッファが所定の状態になった時に、前記第 3 のバッファに対しバックプレッシャ制御を行うことを特徴とする請求項 1 記載の通信装置。

【請求項 3】 前記通信装置は更に、前記スイッチ部に接続され、所定のプロトコルに従った処理を行う処理部を有し、

該処理部は前記第 1 及び第 2 のバッファに接続する第 3 及び第 4 のバッファを有し、

前記第 4 のバッファが所定の状態になった時に、前記第 1 のバッファに対しバックプレッシャ制御を行なうことを特徴とする請求項 1 記載の通信装置。

【請求項 4】 異なるインタフェース間のスイッチング機能を有する通信装置において、

固定長のデータをスイッチングするメインスイッチ部と、該メインスイッチ部の入力及び出力にそれぞれ設けられた第 1 及び第 2 のバッファを回線毎に複数個具備するインタフェース部とを有するスイッチ部を複数有することを特徴とする通信装置。

【請求項 5】 異なるインタフェース間のスイッチングを行なう通信制御方法において、

前記異なるインタフェースで扱われるデータに関連する固定長のデータを一旦

バッファリングしてからスイッチングする段階と、

スイッチングされたデータを一旦バッファリングしてから、回線に向けて送出する段階と

を有する通信制御方法。

【発明の詳細な説明】

【 0 0 0 1 】

【発明の属する技術分野】

本発明はネットワークに接続される通信装置に関し、特に、複数の異なる回線を収容してスイッチングする機能を持ったルータなどの通信装置に関する。より詳細には、本発明はルータ内のキューイングに関する技術、及びトラヒック制御の一態様であるバックプレッシャ制御及びこれに関連する技術に関する。

【 0 0 0 2 】

従来より、複数のネットワークを接続してデータをルーティングして中継するために、ルータと呼ばれる中継装置が用いられている。ルータは、ネットワークのプロトコルの変換やアドレスの変換を行なって、データの中継経路を設定する。

【 0 0 0 3 】

【従来の技術】

図 1 に、ネットワーク構成の一例を示す。ルータ 1 0 ～ 1 3 は異なるネットワークを接続する。例えば、ルータ 1 0 は、Sonet / SDH ADM (Sonet / Synchronous Digital Hierarchy Add-Drop Multiplexer) 1 4、ATM (Asynchronous Transfer Mode) 専用線サービスネットワーク 1 5、OC-4 8 DWDM (Dense Wavelength Division Multiplexing) ネットワーク 1 6、及び OC 4 8 c DWDM ネットワーク 1 7 を相互に接続する。換言すれば、ルータ 1 0 は異なる伝送速度（換言すれば、異なるインタフェース）の回線を収容している。同様に、他のルータ 1 1 ～ 1 3 も異なるネットワークを相互に接続してデータを中継する。

【 0 0 0 4 】

このようなルータは、異なるサイズ（長さ）のパケットを中継処理する。つま

り、扱うパケットは可変長である。また、ある回線が輻輳した時に、その回線へのパケットの流入を防いでパケットの損失を防止するためのバックプレッシャ制御を、イーサネットのポート単位（換言すれば回線単位）に行なう。例えば、あるポートが輻輳した場合には、このポートに接続される回線に対してバックプレッシャ制御を行なう。例えば、ルータがポート毎にバッファを具備している場合には、輻輳したバッファへのパケットの流入を規制する。

【 0 0 0 5 】

【発明が解決しようとする課題】

しかしながら、従来のルータは、異なるネットワークを効率よく収容して効率的な中継処理を行えないという問題点を具備する。具体的には次の通りである。

【 0 0 0 6 】

第1に、従来のルータは可変長パケットを中継処理するため、様々なネットワークを中継しようとする、内部制御が極めて複雑になり、異なる伝送速度毎にQoS (Quality of Service) 制御を行なうことは極めて困難である。その際、バックプレッシャ制御はイーサネットのポート単位に行うため、ATMやPOS (Packet Over Switch) などの異なる伝送速度のパケット処理に対しては、必ずしも効率的なバックプレッシャ制御とはならない。従って、異なる伝送速度毎のQoS制御を効果的にかつ効率的に行なうことはできない。

【 0 0 0 7 】

第2に、出力ポート毎にバッファを持つので、バッファを効率的に利用できない。例えば、ある出力ポートが輻輳している時に他の出力ポートのトラフィックが低い場合には、ルータ全体のバッファの使用効率は低い状態にある。この問題点を解決するために、出力回線（ポート）をアグリゲート（束ねて）して、共通のバッファを設けることが考えられる。しかしながら、ある出力回線のバックプレッシャ制御の要求を受けると、出力回線が輻輳していないデータまでバックプレッシャの影響を受けてしまい、データを出力できないブロッキング状態が発生してしまう。

【 0 0 0 8 】

第3に、通常ルータは二重化されている。これは、信頼性や保守運用の観点に基づく。二重化構成のルータにおいて、運用系と非運用系のいずれからのバックプレッシャでも制御動作を開始するように構成すると、レイテンシーやルータの状態によって運用系、非運用系で状態が不整合となってしまう可能性がある。このため、運用系から非運用系に切り替えを行なった場合に、バックプレッシャで制御された状態やバッファリングの状態によって、データの重複や追い越しが発生してしまう可能性がある。また、非運用系で障害が発生してバックプレッシャを発生したままとなった時に、非運用系の障害であるにもかかわらず、運用系に影響を与えてしまうことになる。

## 【0009】

第4に、ATMの回線はデータのゆらぎや遅延に対して規格がある。ATMの規格を遵守するためには、バックプレッシャ制御を行なわないことが理想であるが、バッファの有効利用の観点からはバックプレッシャ制御を行なうことが望まれる。しかしながら、バックプレッシャ制御を頻繁に行なわなければならない状況では、データのゆらぎや遅延の規格を満足することはできない。

## 【0010】

従って、本発明は上記従来技術の問題点を解決し、異なるネットワークを効率よく収容して効率的に中継処理を行える通信装置及び通信制御方法を提供することを目的とする。

## 【0011】

## 【課題を解決するための手段】

本発明は、異なるインタフェース間のスイッチング機能を有する通信装置において、固定長のデータをスイッチングするメインスイッチ部と、該メインスイッチ部の入力及び出力にそれぞれ設けられた第1及び第2のバッファを具備するインタフェース部とを設けたスイッチ部を有することを特徴とする通信装置である。

## 【0012】

固定長のデータをスイッチングするメインスイッチ部の入力及び出力にそれぞれ、第1及び第2のバッファを設けたことで、メインスイッチ部をバッファレス



な構成とすることができる。従って、ネットワークのプロトコルに依存した伝送速度差を吸収することができ、スイッチングによるゆらぎを低減することができる。そして、伝送速度毎のQoS制御やバックプレッシャ制御を効果的かつ効率的に行なうことができる。

【0013】

【発明の実施の形態】

図2は、本発明の第1の実施の形態によるルータの構成を示すブロック図である。図示するルータ20は、図1のルータ10～13に置き換わるものである。図示するルータ20は二重化されている。

【0014】

ルータ20は、回線終端部21w、21p、L3（レイヤ3：ネットワーク層）処理部22w、22p、スイッチ部23w、23p、L3処理部24w、24p及び回線終端部25w、25pを有する。添え字の“w”は二重化構成の運用系を示し、“p”は非運用系を示す。スイッチ部23w、23pに接続される二重化構成のL3処理部や回線終端部は、必ずしも二重化にする必要はない。

【0015】

ルータ20が扱う回線の種類は例えば、Ethernet 10/100Base-T、Ethernet 1000Base-T、POS OC3C、OC12C、OC48C、ATM OC3C、OC12C、STM T1/E1、STS3、STS12などである。

図2の各部の内部構成例

図3は、図2の各部の内部構成例を示すブロック図である。図3に示す構成は、二重化構成の運用系及び非運用系のいずれにも当てはまるものである。よって、図3では、運用系と非運用系を区別する参照番号の添え字“w”や“p”を省略してある。また、図3に示す構成は実線の矢印で示すデータの流れに沿って図示したものである。従って、回線終端部21及びL3処理部22は、入力側と出力側とで2分割されたように図示してある。

【0016】

入力側の回線終端部21は、物理層処理部211及びL3インタフェース部2

1 2 を有する。L 3 処理部 2 2 は、回線インタフェース部 2 2 1、ローカルスイッチ部 2 2 2 及びスイッチインタフェース部 2 2 3 を有する。L 3 処理部 2 2 は、複数のネットワークに接続されている通信装置間のデータ転送やデータの中継処理などの通信プロトコルに従った処理を行なう。スイッチ部 2 3 は、入力側のスイッチインタフェース部 2 3 1、メインスイッチ部 2 3 2 及び出力側のスイッチインタフェース部 2 3 3 を有する。出力側の L 3 処理部 2 2 は、スイッチインタフェース部 2 2 4、ローカルスイッチ部 2 2 5 及び回線インタフェース部 2 2 6 を有する。出力側の回線終端部 2 5 は、L 3 インタフェース部 2 1 3 及び物理層処理部 2 1 4 を有する。

## 【 0 0 1 7 】

回線終端部 2 1 の物理層処理部 2 1 1 は、ポートを介して接続されるネットワーク上の回線を収容してアグリゲートする。L 3 インタフェース部 2 1 2 は、物理層処理部 2 1 1 でアグリゲートした回線上のデータに対し、レイヤ 2 の処理を行なう（レイヤ 2 終端処理）。L 3 処理部 2 2 の回線インタフェース部 2 2 1 は、可変長パケットを一旦バッファに蓄積した後、所定長の固定長パケット（以下セルと言う）に変換する。この処理をフラグメント化という。ローカルスイッチ部 2 2 2 は、回線インタフェース部 2 2 1 からのセルをスイッチングする。スイッチインタフェース部 2 2 3 は、ローカルスイッチ部 2 2 2 が出力するセルを一旦蓄積した後、スイッチ部 2 3 に出力する。スイッチ部 2 3 は、L 3 処理部 2 2 からのセルを一旦蓄積した後、メインスイッチ部 2 3 2 にセルを出力する。メインスイッチ部 2 3 2 は、レイヤ 3 の IP ( Internet Protocol ) に基づきルーティングを行なう。メインスイッチ部 2 3 2 はバッファを持たない。

## 【 0 0 1 8 】

図 4 に、レイヤ 2 の終端処理とレイヤ 3 のルーティング処理を示す。例えば、ATM の回線では、ATM レイヤは回線終端部 2 1 で終端しておき、レイヤ 3 の IP 情報を元にスイッチ部 2 3 が ATM セルをルーティングする。

## 【 0 0 1 9 】

スイッチインタフェース部 2 3 3 は、メインスイッチ部 2 3 2 でルーティング

されたセルを一旦格納する。

#### 【0020】

出力側のL3処理部22のスイッチインタフェース部224は、スイッチ部23からのセルを一旦格納する。ローカルスイッチ部225は、回線インタフェース部224からのセルをスイッチングする。回線インタフェース部226は、ローカルスイッチ部225からのセルを一旦格納する。回線終端部21のL3インタフェース部213は、L3処理部22からのルーティングされたセルを一旦格納してレイヤ2に関する情報を付加し、更にセルに対応する可変長パケットに変換する。物理層処理部214は、可変長パケットに対応する回線（ポート）に出力する。

#### 【0021】

図3中の記号“×”は、後述するバックプレッシャ信号の廃棄（終端）ポイントを示す。本発明の第1の実施の形態では、L3処理部22のスイッチインタフェース部223でバックプレッシャ信号を終端し、データを廃棄する。また、後述する本発明の第2の実施の形態では、L3処理部22のローカルスイッチ部225でバックプレッシャ信号を終端し、データを廃棄する。

#### L3処理部22及びスイッチ部23の基本構成及び基本動作

図5は、L3処理部22及びスイッチ部23の基本構成を示すブロック図である。L3処理部22は回線対応処理部22<sub>0</sub>～22<sub>7</sub>及び内部処理部22<sub>8</sub>を有する。回線対応処理部22<sub>0</sub>は、インタフェース部220<sub>0</sub>及びネットワークプロセッサ227<sub>0</sub>を有する。インタフェース部220<sub>0</sub>はルータ20の1つのポート（回線）に対応し、図3のスイッチインタフェース部223及び224の内部回路に相当する。ネットワークプロセッサ227<sub>0</sub>は、図3のローカルスイッチ部222に接続されるもので、図面を簡単にするために、図3での図示を省略してある。ネットワークプロセッサ227<sub>0</sub>は、入力2系統IN0、IN1と出力2系統OUT0、OUT1を処理する。インタフェース部220<sub>0</sub>は、FIFO形式の共通バッファ31及び32、入力バッファ35、並びに出力バッファ36を具備する。他の回線対応処理部22<sub>2</sub>～22<sub>7</sub>も同様の構成である。すなわち、回線対応処理部22<sub>2</sub>～22<sub>7</sub>はそれぞれ、インタフェース部220<sub>1</sub>～

220<sub>7</sub>及びネットワークプロセッサ227<sub>1</sub>～227<sub>7</sub>を有する。

【0022】

内部処理部22<sub>8</sub>は、インタフェース部220<sub>8</sub>とメインプロセッサ227<sub>8</sub>とを有する。メインプロセッサ227<sub>8</sub>は、ルータ20を統合的に制御するプロセッサであり、インタフェース部220<sub>8</sub>を介してスイッチ部23との間でセルをやり取りする。

【0023】

スイッチ部23は、図3に示すメインスイッチ232と回線対応処理部230<sub>0</sub>～230<sub>7</sub>とを有する。インタフェース部230<sub>0</sub>～230<sub>7</sub>は、図3のスイッチインタフェース231、233の内部回路に相当する。インタフェース部230<sub>0</sub>～230<sub>7</sub>はそれぞれ、L3処理部のインタフェース部220<sub>0</sub>～220<sub>7</sub>に対応している。また、インタフェース部220<sub>8</sub>はインタフェース部230<sub>7</sub>に接続されている。インタフェース部230<sub>0</sub>は、FIFO形式の共通バッファ33及び34、出力バッファ37並びに入力バッファ38を有する。

【0024】

他のインタフェース部230<sub>1</sub>～230<sub>7</sub>も同様に構成されている。

【0025】

なお、便宜上、共通バッファ33と出力バッファ37とをまとめて第1のバッファと言う。また、共通バッファ34と入力バッファ38とをまとめて第2のバッファと言う。更に、共通バッファ35と入力バッファ31とをまとめて第3のバッファと言う。更に、共通バッファ32と出力バッファ36とをまとめて第4のバッファと言う。

【0026】

次に、図5に示す回路の基本動作を説明する。

【0027】

ネットワークプロセッサ227<sub>0</sub>は、ローカルスイッチ222からセルを受け取り、入力バッファ35に出力する。セルの入力は、入力IN0とIN1の2系統ある。入力バッファ35は受取ったセルを共通バッファ31に出力する。これにより、共通バッファ31にはセルのキューが形成される。共通バッファ31内

のセルは、後述するスケジューリング処理に従い読み出され、スイッチ部 2 3 に送られる。

#### 【 0 0 2 8 】

スイッチ部 2 3 の共通バッファ 3 3 は、インタフェース部 2 2 0<sub>0</sub> から送られたセルを格納する。共通バッファ 3 3 に格納されたセルは、後述するスケジューリング処理に従い読み出され、出力バッファ 3 7 に一旦格納された後、メインスイッチ部 2 3 2 に送られる。メインスイッチ部 2 3 2 は、受取ったセルをスイッチングする。

#### 【 0 0 2 9 】

メインスイッチ部 2 3 2 からのセルは、入力バッファ 3 8 に一旦格納された後、共通バッファ 3 4 に格納される。共通バッファ 3 4 に格納されたセルは、後述するスケジューリング処理に従い読み出され、インタフェース部 2 2 0<sub>0</sub> に送られる。インタフェース部 2 2 0<sub>0</sub> の共通バッファ 3 2 は、受取ったセルを格納する。そして、後述するスケジューリング処理に従いセルが共通バッファ 3 2 から読み出され、一旦出力バッファ 3 6 に格納される。そして、出力バッファ 3 6 から読み出されたセルは、ネットワークプロセッサ 2 2 7<sub>0</sub> に出力される。ネットワークプロセッサ 2 2 7<sub>0</sub> は、出力系 O U T 0 又は O U T 1 を介して、受取ったセルをローカルスイッチ部 2 2 2 に出力する。

#### 【 0 0 3 0 】

後述するように、入力バッファ 3 5、3 8 及び出力バッファ 3 6、3 7 の各々は、Q o S クラス毎にバッファ(キュー)を有している。Q o S クラスは例えば、固定ビットレートのサービス、可変ビットレートのサービス、無制限ビットレートのサービス、アベイラブル ( a v a i l a b l e ) ビットレートのサービスや、マルチキャストのサービスなどがある。Q o S サービス単位はインタフェース毎に任意に設定可能である。

#### バックプレッシャ制御の概要

図 5 には、本発明の第 1 の実施の形態で用いられるバックプレッシャ信号 B P 1、B P 2、B P 3、B P 4 及び B P 5 を示している。バックプレッシャ信号 B P 1 は、スイッチ部 2 3 のインタフェース部 2 3 0<sub>0</sub> 内の共通バッファ 3 3 又は

出力バッファ 3 7 が所定の状態になると生成される。所定の状態とは例えば、輻射した状態又は輻射が予期される状態である。輻射が予期される状態を含めて、輻射と定義しても良い。

#### 【 0 0 3 1 】

バックプレッシャ信号 B P 1 は、L 3 処理部 2 2 のインタフェース部 2 2 0 0 の内部に設けられた共通バッファ 3 1 からのセル読み出し動作を停止させる。図 5 に示すバックプレッシャ信号 B P 1 の矢印は論理的な流れを示し、好ましくは、バックプレッシャ信号 B P 1 をセルで構成する。このセルをフロー制御セルと言う。つまり、インバンドフロー制御でバックプレッシャ信号 B P 1 を伝送する。具体的には、共通バッファ 3 3 又は出力バッファ 3 7 が所定の状態になると、スケジューリング処理が共通バッファ 3 4 を介してフロー制御セルをインタフェース部 2 2 0 0 の共通バッファ 3 2 に送る。共通バッファ 3 2 からフロー制御セルが読み出されると、スケジューリング処理は共通バッファ 3 1 のセル読み出し動作を停止させる。

#### 【 0 0 3 2 】

バックプレッシャ信号 B P 2 は、L 3 処理部 2 2 のインタフェース部 2 2 0 0 内部の共通バッファ 3 2 又は出力バッファ 3 6 が所定の状態になると生成される。バックプレッシャ信号 B P 2 は、スイッチ部 2 3 のインタフェース部 2 3 0 0 内部に設けられた共通バッファ 3 3 からのセル読み出し動作を停止させる。図 5 に示すバックプレッシャ信号 B P 2 の矢印は論理的な流れを示し、好ましくは、バックプレッシャ信号 B P 2 をフロー制御セルで構成する。フロー制御セルの物理的な流れは次の通りである。共通バッファ 3 2 又は出力バッファ 3 6 が所定の状態になると、スケジューリング処理が共通バッファ 3 1 を介してフロー制御セルをインタフェース部 2 3 0 0 の共通バッファ 3 3 に送る。共通バッファ 3 3 からフロー制御セルが読み出されると、スケジューリング処理は共通バッファ 3 3 のセル読み出し動作を停止させる。

#### 【 0 0 3 3 】

なお、後述するように、バックプレッシャ信号 B P 2 はリンクレベルのフロー制御、つまりインタフェース部 2 3 0 0 ~ 2 3 0 7 に設けられた全ての共通バッ

ファ 33 の読み出しを停止させることもできる。この制御は、後述するバックプレッシャ信号 BP5 と協働して行なわれる。

#### 【0034】

バックプレッシャ信号 BP3 は、スイッチ部 23 のインタフェース部 230<sub>0</sub> 内部の共通バッファ 34 又は入力バッファ 38 が所定の状態になると生成される。バックプレッシャ信号 BP3 は、スイッチ部 23 のインタフェース部 230<sub>0</sub> 内部に設けられた共通バッファ 33 からのセル読み出し動作を停止させる。図 5 に示すバックプレッシャ信号 BP3 の矢印は論理的な流れを示し、好ましくは、バックプレッシャ信号 BP2 をフロー制御セルで構成する。具体的には、共通バッファ 34 又は入力バッファ 38 が所定の状態になると、スケジューリング処理が共通バッファ 31 を介してフロー制御セルをインタフェース部 230<sub>0</sub> の共通バッファ 33 に送る。共通バッファ 33 からフロー制御セルが読み出されると、スケジューリング処理は共通バッファ 33 のセル読み出し動作を停止させる。

#### 【0035】

バックプレッシャ信号 BP4 は、出力 OUT0、OUT1 単位に共通バッファ 32 のセル読み出しを制御する。出力 OUT0 又は OUT1 の先にあるローカルスイッチ 225（図 3 参照）などに設けられている内部バッファが所定の状態になったとき、ネットワークプロセッサ 227<sub>0</sub> は専用線を介して共通バッファ 32 にバックプレッシャ信号 BP4 を送出する。

#### 【0036】

バックプレッシャ信号 BP5 は、後述するバックプレッシャ・バスを介してシリアルに伝送される信号である。バックプレッシャ・バスは、インタフェース部 230<sub>0</sub>～230<sub>7</sub> を相互に接続する。バックプレッシャ信号 BP5 は、インタフェース部 230<sub>0</sub>～230<sub>7</sub> 内部の共通バッファ 33 のセル読み出しを同時に停止させる。バックプレッシャ信号 BP5 による読み出しは、後述するように、QoS クラス単位(サービスクラスとも言う)及びバッファ単位に停止させることができる。

インタフェース部 220<sub>0</sub>～220<sub>7</sub> 内に設けられた共通バッファ 31 のスケジューリング処理

インタフェース部 220<sub>0</sub>～220<sub>7</sub>の共通バッファ 31 のスケジューリング処理について、図 6 を参照して説明する。

【0037】

図 6 は、このスケジューリング処理を行うスケジューラの一構成例を示す図である。図 6 のスケジューラ（以下、第 1 のスケジューラと言う）は、インタフェース部 220<sub>0</sub>～220<sub>7</sub>の共通バッファ 31 に対応したアドレスキュー部 41<sub>0</sub>～41<sub>7</sub>、マルチキャストすべきセルのアドレスを格納するアドレスキュー部 45、アドレスキュー部 41<sub>0</sub>～41<sub>7</sub>の出力を選択するセレクタ 44、及びセレクタ 44 の出力及びアドレスキュー部 45 の出力を選択するセレクタ 45 とを具備する。前述した図 1 から図 5 までの図には、便宜上第 1 のスケジューラの図示を省略してある。上記第 1 のスケジューラは、図 5 の①の部分の読み出しを制御する。

【0038】

アドレスキュー部 41<sub>0</sub>～41<sub>7</sub>の各々は、クラスに対応した数のキュー 43<sub>0</sub>～43<sub>7</sub>を有する。この例では、クラス 0 からクラス 7 までの 8 つのクラスがある場合を想定している。これらのキュー 43<sub>0</sub>～43<sub>7</sub>に、対応する共通バッファ 31 に格納されたセルのアドレスポインタ値が格納される。例えば、クラス 0 のセルが共通バッファ 31 に格納された場合には、このセルのアドレスポインタ値がアドレスキュー部 41<sub>0</sub>のキュー 43<sub>0</sub>に格納される。各キュー 43<sub>0</sub>～43<sub>7</sub>は F I F O 形式のメモリで構成される。アドレスキュー部 45 には、マルチキャストすべきキューが格納された共通バッファ 31 のアドレスポインタ値が格納される。

【0039】

セレクタ 42 は、クラス間のスケジューリングに従い、読み出すべきキューを選択する（調停する）。このスケジューリングの選択論理は、重み付けラウンド・ロビン方式（Weighted Round Robin：以下、WRR 方式と言う）とする。通常のラウンド・ロビン方式（RR 方式）が単純な巡回選択であるのに対し、WRR 方式は巡回するキューに重み付けを設定することが可能である。この重み付けは、そのキューから連続で読み出しを行なうことができる最



高回数を規定するもので、全てのキューの重み付けを1とした場合には、WRR方式はRR方式と同一になる。初期化後、キュー43<sub>0</sub>から選択を行ない、キューがエンプティ又は連続読み出しを行なったという条件で、次のパケット時間において次のクラスの読み出しへと移行する。重み付けは、各クラス毎に設定できる。例えば、インタフェース部220<sub>0</sub>～220<sub>7</sub>に共通としても良い。

## 【0040】

このようにして、アドレスキュー部41<sub>0</sub>～41<sub>7</sub>の各々でクラス間のスケジューリングが行なわれる。

## 【0041】

セクタ44は、アドレスキュー部41<sub>0</sub>～41<sub>7</sub>の中から、セルを読み出すべきアドレスキュー部を選択するスケジューリング処理を行う。選択論理は、例えばRR方式である。

## 【0042】

セクタ46は、セクタ44の出力とマルチキャストのアドレスキュー部45の出力とのいずれか一方を選択する。この選択論理は、セクタ44からの出力が無い時、換言すれば、アドレスキュー部41<sub>0</sub>～41<sub>7</sub>の各共通バッファ31から読み出すべきユニキャストのセルが無い場合に限り、マルチキャストの読み出しを行なう。この場合には、共通バッファ31内のマルチキャストキュー（バッファ）が読み出し対象となる。マルチキャストキューにも読み出すべきセルが存在しない場合には、そのセル時間はセルを読み出さない。

## 【0043】

このようにして、第1のスケジューラはインタフェース部220<sub>0</sub>～220<sub>7</sub>から読み出すべきセルのアドレスを決定する。

## 【0044】

なお、第1のスケジューラは後述するバックプレッシャ信号BP2を形成するフロー制御セルを送出するために、強制的に1セルだけ無効セルを挿入し、バッファからのセル読み出しを行なわない時間を作ることができる。

### バックプレッシャ信号BP1によるバックプレッシャ制御

以上のようにして、3通りのスケジューリングにより任意のパケット時間の読

み出しキューを決定する。その際、バックプレッシャ信号BP1が上記スケジューラに通知されると、第1のスケジューラはバックプレッシャ信号BP1に従いセル読み出し処理を停止させる。以下に説明するように、バックプレッシャ信号BP1には2通りある。

#### 【0045】

バックプレッシャ信号BP1がリンクレベルの場合、つまりスイッチインタフェース部23がスイッチインタフェース部22に対し全てのセルの読み出しの停止を要求する場合には、第1のスケジューラは図5に示すバックプレッシャ信号BP1を受け取り、図6に示すセクタ46をディスエーブル状態とする。従って、セル読み出しのアドレスが共通バッファ31に供給されず、インタフェース部220<sub>0</sub>～220<sub>7</sub>からのセルの読み出しは停止する。

#### 【0046】

これに対し、バックプレッシャ信号BP1がポート単位（回線単位）の場合、例えばスイッチインタフェース部23のインタフェース部230<sub>0</sub>内部の共通バッファ33が輻輳してしまう可能性がある。よって、共通バッファ31からのセル読み出しの停止を要求する場合には、第1のスケジューラは図5に示すバックプレッシャ信号BP1を受け取り、図6に示すセクタ42をディスエーブル状態とする。従って、インタフェース部220<sub>0</sub>の共通バッファ31からのセル読み出しは停止する。

インタフェース部220<sub>0</sub>～220<sub>7</sub>の内部に設けられた共通バッファ32のスケジューリング処理

インタフェース部220<sub>0</sub>～220<sub>7</sub>の共通バッファ32のスケジューリング処理は、第2のスケジューラが行なう。第2のスケジューラは、図5の②の部分のセル読み出しを制御する。

#### 【0047】

図7は、第2のスケジューラの一構成例を示す図である。第2のスケジューラは、出力OUT0とOUT1にそれぞれ対応したアドレスキュー部47<sub>0</sub>、47<sub>1</sub>と、アドレスキュー部47<sub>0</sub>、47<sub>1</sub>の出力を選択するセクタ51とを具備する。各アドレスキュー部47<sub>0</sub>、47<sub>1</sub>は、8つのクラスに対応するキュー4

8<sub>0</sub> ~ 4 8<sub>7</sub>、マルチキャストに対応するキュー 4 8<sub>8</sub>、これらのキューの 1 つを選択するセクタ 4 9、及びセクタ 4 9 とキュー 4 8<sub>8</sub> のいずれかを選択するセクタ 5 0 とを有する。

#### 【 0 0 4 8 】

第 2 のスケジューラは、各パケット時間毎に、どの Q o S クラスに対応するキューを読み出すかを決定する。このための選択処理には、以下の 2 つがある。第 1 の論理は出力 O U T 0、O U T 1 間の調停であり、第 2 の論理は Q o S クラス間の調停である。

#### 【 0 0 4 9 】

第 1 の論理は、各パケット時間毎に O U T 0 と O U T 1 のどちら宛のパケットを読み出すかを決定する。この選択は、パケット時間毎に固定的に交互に O U T 0 と O U T 1 を読出すものである。このために、2 パケット時間毎のマルチフレームを生成し、この前半で O U T 0、後半で O U T 1 の読み出しを行う。

#### 【 0 0 5 0 】

第 2 の論理は、各出力 O U T 0、O U T 1 毎に、8 つのクラス対応のキュー 4 8<sub>0</sub> ~ 4 8<sub>7</sub>、及び 1 つのマルチキャスト用キュー 4 8<sub>8</sub> の中から読み出すキューを決定する。選択論理の一例として、まずフレームの連続性を保証する論理を最優先とし、次にユニキャストの 8 つの Q o S についての W R R 方式を用い、最後にユニキャストとマルチキャストとの間の固定優先度に従った選択とする。フレームの連続性を保証する論理では、選択中のキューから連続してセルを読み出すことができるように、アドレスポインタ値を対応するキューから読み出す。いずれのキューもフレーム読み出し状態にない場合に、次の W R R 方式に移る。この W R R 方式を用いた論理選択は、前述した共通バッファ 3 1 のスケジューリング処理で説明した選択と同様である。ただし、フレームの連続性を保証する必要があるため、重み付けはセルの連続読み出し回数ではなく、フレームの連続読み出しの最高回数となる。固定優先度に従った処理では、全てのクラスのキュー 4 8<sub>0</sub> ~ 4 8<sub>7</sub> がエンプティの場合には、マルチキャストに対応するキュー 4 8<sub>8</sub> からアドレスポインタ値を読み出す。つまり、マルチキャストのフレームはユニキャストのフレームが存在する限り読み出されることはない。なお、全てのキュー

48<sub>0</sub>～48<sub>8</sub> がエンプティ状態の場合には、読み出しは無効である。

#### バックプレッシャ信号BP4によるバックプレッシャ制御

前述したように、バックプレッシャ信号BP4は、出力OUT0、OUT1単位に共通バッファ32のセル読み出しを制御する。出力OUT0又はOUT1の先にあるローカルスイッチ225（図3参照）などに設けられている内部バッファが所定の状態になったとき、ネットワークプロセッサ227<sub>0</sub>は専用線を介して共通バッファ32にバックプレッシャ信号BP4を送出する。バックプレッシャ信号BP4を受けた図7に示す第2のスケジューラは、バックプレッシャ信号BP4で指定された出力に対応するアドレスキュー部47<sub>0</sub>、47<sub>1</sub>の出力をマスクする。マスクすることで、読み出すべきパケットが存在しないものとする。つまり、アドレスキュー部47<sub>0</sub>、47<sub>1</sub>の出力を無効にする。

#### インタフェース部230<sub>0</sub>～230<sub>7</sub>の内部に設けられた共通バッファ33のスケジューリング処理

インタフェース部230<sub>0</sub>～230<sub>7</sub>の共通バッファ33のスケジューリング処理について、図8及び図9を参照して説明する。

##### 【0051】

図8は、インタフェース部230<sub>0</sub>の共通バッファ33及び出力バッファ37の一構成例を示す図である。共通バッファ33は、前処理部53、共通バッファ33、アドレス制御部55、及びスケジューラ70を有する。スケジューラ（以下、第3のスケジューラと言う）は、キュー振り分け部56、アドレスキュー部57<sub>0</sub>、57<sub>1</sub>、及びセクタ59を有する。出力バッファ37は、セル分配部66、FIFO形式のバッファ61、62、セクタ63<sub>0</sub>、63<sub>1</sub>、及びFIFO形式のバッファ64<sub>0</sub>、64<sub>1</sub>を有する。第3のスケジューラは、図5の③の部分のセル読み出しを制御する。

##### 【0052】

前処理部53は、L3処理部22の回線対応処理部22<sub>0</sub>から受取ったセルを内部クロックに乗せ換えた後、固定長正規化処理を行う。セルは、アドレス制御部55が発行する格納アドレスに従い、共通バッファ33に格納される。この格納アドレスは例えば、アドレス制御部55の書き込みアドレス発行機能がシーケ

ンシャルに発行する。各セルの宛先情報、発信元情報及びQoSなどの各情報は、アドレス制御部55を経由してキュー振り分け部56に出力される。

#### 【0053】

キュー振り分け部56は、各セルに付与されている宛先情報、発信元情報及びQoSと上記格納アドレスをアドレス制御部55から受け取り、書き込みアドレス、つまりアドレスポインタ値をアドレスキュー部57<sub>0</sub>又は57<sub>1</sub>の内部キューに書き込む。

#### 【0054】

アドレスキュー部57<sub>0</sub>は入力IN0に対応し、内部に、インタフェース部230<sub>0</sub>～230<sub>7</sub>に対応するキュー58<sub>0</sub>～58<sub>7</sub>、図5のインタフェース部228<sub>8</sub>に対応するキュー58<sub>8</sub>、及びマルチキャストに対応するキュー58<sub>8</sub>を具備している。同様に、アドレスキュー部57<sub>1</sub>は入力IN1に対応し、内部に、インタフェース部230<sub>0</sub>～230<sub>7</sub>に対応するキュー58<sub>0</sub>～58<sub>7</sub>、図5のインタフェース部228<sub>8</sub>に対応するキュー58<sub>8</sub>、及びマルチキャストに対応するキュー58<sub>8</sub>を具備している。セクタ59は、アドレスキュー部57<sub>0</sub>と57<sub>1</sub>の出力を選択する。

#### 【0055】

図9は、第3のスケジューラ70の一構成例を示す図である。スケジューラ70は、入力IN0、IN1内部のスケジューリング動作及び入力IN0、IN1間のスケジューリング動作を行う。第3のスケジューラ70は、入力IN0に対応してセクタ65、66及び67を有する。図示を省略してあるが、第3のスケジューラ70は、入力IN1に対応して同様のセクタを有する。

#### 【0056】

セクタ65は、アドレスキュー58<sub>7</sub>と58<sub>8</sub>との間のスケジューリングを行なう。つまり、このスケジューリングは、アドレスキュー58<sub>0</sub>～58<sub>6</sub>間のスケジューリングの前段で行なわれる。このスケジューリングは例えば、RR方式で行なわれる。セクタ65で選択されたアドレスポインタ値は、アドレスキュー58<sub>0</sub>～58<sub>6</sub>とともにセクタ66でスケジューリング処理される。このスケジューリングは例えば、RR方式で行なわれる。セクタ65で選択された

アドレスポインタ値は、セクタ 6 7 に送出される。セクタ 6 7 は、セクタ 6 5 の出力、つまりユニキャストのセルと、アドレスキュー 5 8<sub>g</sub> に格納されているアドレスポインタで指定されるマルチキャストのセルとの間のスケジューリングを行なう。このスケジューリングは、ユニキャストキューが存在しない場合にのみ、マルチキャストキューからの読み出しを可能とする。

## 【 0 0 5 7 】

最後に、アドレスキュー 5 7<sub>0</sub> と 5 7<sub>1</sub> のセクタ 6 7 の出力をセクタ 5 9 内部の図示を省略するセクタで選択し（例えば交互に選択する）、選択したアドレスポインタ値をアドレス制御部 5 5 に出力する。アドレス制御部 5 5 の読み出しアドレス発行機能は、スケジューラ 7 0 から受取ったアドレス値に基づき、読み出しアドレスを発行して共通バッファ 3 3 に出力する。

## 【 0 0 5 8 】

共通バッファ 3 3 から読み出されたセルは、出力バッファ 3 7 のキュー振り分け部 6 0 で、入力 IN 0 に対応するバッファ 6 1 又は入力 IN 1 に対応するバッファ 6 2 に振り分ける。セクタ 6 3<sub>0</sub> と 6 3<sub>1</sub> はそれぞれ、バッファ 6 1 と 6 2 から読み出されたセルを選択して、図 5 に示すメインスイッチ部 2 3 2 に出力する。

バックプレッシャ信号 B P 5 によるバックプレッシャ制御

図 1 0 は、バックプレッシャ信号 B P 5 によるバックプレッシャ制御を説明するための図である。前述した図に記載の構成要素と同一のものには、同一の参照番号を付してある。また、図 1 0 は、図 8 を参照して説明したインタフェース部 2 3 0<sub>0</sub> の内部構成を概略的に示している。同様に、インタフェース部 2 3 0<sub>7</sub> の内部構成も示している。

## 【 0 0 5 9 】

前述したように、バックプレッシャ信号 B P 5 は、バックプレッシャ・バスを介してシリアルに伝送される信号である。このバックプレッシャ・バスを、図 1 0 において参照番号 7 1 で示す。バックプレッシャ・バス 7 1 は、インタフェース部 2 3 0<sub>0</sub> ~ 2 3 0<sub>7</sub> を相互に接続する。

## 【 0 0 6 0 】

図 1 0 に示す “R D Y” は、フロー制御セルによりインタフェース部 2 2 0<sub>0</sub> から通知されたバックプレッシャ信号 B P 2 を意味する。このバックプレッシャ信号 B P 2 は前述したポート単位の要求の他、リンクレベルのバックプレッシャを要求も含む。バックプレッシャ B P 2 が所定の値の時、バックプレッシャ B P 2 はクラス単位でかつ全てのインタフェース部 2 3 0<sub>0</sub> ~ 2 3 0<sub>7</sub> の対応するバッファからの読み出しの停止（リンクレベルのバックプレッシャ）を要求する。例えば、図 1 0 に示す例では、クラス j（j はクラス 0 から 7 のいずれか）とマルチキャストに関してセル読み出しの停止が求められている。

## 【 0 0 6 1 】

このようなバックプレッシャ信号 B P 2 を受けた第 3 のスケジューラは、自己インタフェース 2 3 0<sub>0</sub> のキュー 5 8<sub>j</sub> と、マルチキャストに対応するキュー 5 8<sub>g</sub> を制御して、これらのバッファからのセル読み出しを停止させる。具体的には、図 9 に示すセクタ 6 6 と 6 7 の選択論理において、キュー 5 8<sub>j</sub> と 5 8<sub>g</sub> を選択しないようにする。同時に、第 3 のスケジューラは、バックプレッシャ・バス 7 1 を介してバックプレッシャ信号 B P 5 を他のインタフェース部 2 3 0<sub>1</sub> ~ 2 3 0<sub>7</sub> に出力する。このバックプレッシャ信号 B P 5 は、キュー 5 8<sub>j</sub> と 5 8<sub>g</sub> を指定する情報を含む。図 1 0 には、バックプレッシャ信号 B P 5 がインタフェース部 2 3 0<sub>7</sub> のキュー 5 8<sub>j</sub> と 5 8<sub>g</sub> のセル読み出し停止を指示している様子を示している。

## 【 0 0 6 2 】

上記バックプレッシャの制御により、全てのインタフェース部 2 3 0<sub>0</sub> ~ 2 3 0<sub>7</sub> からのインタフェース部 2 2 0<sub>0</sub> 宛セルの送出を停止させることができる。  
バックプレッシャ信号 B P 3 によるバックプレッシャ制御

前述したように、バックプレッシャ信号 B P 3 は、スイッチ部 2 3 のインタフェース部 2 3 0<sub>0</sub> 内部の共通バッファ 3 4 又は入力バッファ 3 8 が所定の状態になると生成される。

## 【 0 0 6 3 】

図 1 1 は、インタフェース部 2 3 0<sub>0</sub> の入力バッファ 3 8 と共通バッファ 3 4 を模式的に図示してある。図 1 1 の構成は、出力 O U T 0 と O U T 1 に対応する

バッファ 6 9<sub>0</sub>、6 9<sub>1</sub> とセレクタ 7 2 とを有する。バッファ 6 9<sub>0</sub> 又は 6 9<sub>1</sub> の使用量が所定のしきい値を超えた場合には、メインスイッチ部 2 3 2 へバックプレッシャ信号 B P 3 を出力する。前述したように、バックプレッシャ信号 B P 3 は、フロー制御セルで構成される。バッファ 6 9<sub>0</sub> 又は 6 9<sub>1</sub> にセルが滞留するのは、風呂ー制御セルが集中した場合のみであり、通常ではメインスイッチ部 2 3 2 からの受信レートよりもバッファ 6 9<sub>0</sub>、6 9<sub>1</sub> からの送信レートの方が高い。よって、バッファ 6 9<sub>0</sub>、6 9<sub>1</sub> にはセルが溜まり難く、フロー制御セルのトラヒックが集中しない限り、バックプレッシャ信号 B P 3 がイネーブルになることはない。

#### 【 0 0 6 4 】

バックプレッシャ信号 B P 3 を受けたメインスイッチ部 2 3 2 は、図 1 2 に示すように、デマルチプレクサ 3 7 の対応する入力（図 1 2 の例では、入力 I N 0）のバッファ 6 4<sub>0</sub> に出力する。この結果、バッファ 6 4<sub>0</sub> のセル読み出しは停止される。

#### バックプレッシャ信号 B P 1 の生成

図 1 3 に、バックプレッシャ信号 B P 1 の生成の一例を示す。図示する例は、インタフェース部 2 3 0<sub>0</sub> における入力 I N 0 の出力バッファ 3 7 を構成するバッファ 6 1 内部のクラス j に対応するキュー 5 8<sub>j</sub> が輻輳した場合又は輻輳する可能性がある場合に、バックプレッシャ信号 B P 1 が生成される。前述したように、バックプレッシャ B P 1 は、フロー制御セルで形成される。

#### 【 0 0 6 5 】

インタフェース部 2 3 0<sub>0</sub> は、セル滞留カウンタ 7 5 を具備している。セル滞留カウンタ 7 5 は、共通バッファ 3 3 に格納される各クラスのセル、図 5 に示すメインプロセッサ 2 2 7<sub>8</sub> 宛のセル、及びマルチキャストされるセルをそれぞれカウントする。図 1 3 に示すセル滞留カウンタ 7 5 内部のキューを示す参照番号 5 8<sub>0</sub> ～ 5 8<sub>9</sub> は、それぞれのカウンタを示す。カウント値は例えば、流入したセルの総数、又は所定時間内で流入したセルの数（セルの流量）を示す。流入したセルの総数又は流量に対するしきい値が設定されている。セル滞留カウンタ 7 5 内部に設けられた比較器は、カウント値と所定のしきい値とを比較する。比較



結果がしきい値超え又は流量超過を示している場合には、バックプレッシャ信号 B P 1 を形成するフロー制御セルが出力 O U T 1 に送られる。このフロー制御セルは、キュー 5 8<sub>j</sub> を識別する情報やユニキャスト・マルチキャストを識別する情報を含む。

#### バックプレッシャ信号 B P 5 の生成

図 1 4 に、バックプレッシャ信号 B P 5 の生成の一例を示す。図示する例は、インタフェース部 2 3 0<sub>0</sub> における入力 I N 0 の共通バッファ 3 3 に流入するセルのしきい値超え又は流量超過が発生した場合には、バックプレッシャ信号 B P 2 に相当する前述の R D Y 信号をフロー制御セルでインタフェース部 2 2 0<sub>0</sub> に送出する。

#### バックプレッシャ信号 B P 5 によるバックプレッシャ制御の他の例

図 1 0 を参照して説明したバックプレッシャ信号 B P 5 によるバックプレッシャ制御は、全てのインタフェース部 2 3 0<sub>0</sub> ~ 2 3 0<sub>7</sub> からのインタフェース部 2 2 0<sub>0</sub> 宛セルの送出を停止させるものであった（リンクレベルの制御）。これに対し、図 1 5 を参照して以下に説明するバックプレッシャ制御は、クラス単位の制御である。

#### 【 0 0 6 6 】

図 1 5 は、L 3 処理部 2 2 の回線対応処理部 2 2<sub>7</sub> 内部のインタフェース部 2 2 0<sub>7</sub> に設けられた出力バッファ 3 6 内のあるクラスに対応するキュー内のセル数又は流量が所定のしきい値を超えた場合を示す。この場合、バックプレッシャ信号 B P 2 を形成するフロー制御セル F C がインタフェース部 2 2 0<sub>7</sub> から対応するインタフェース部 2 3 0<sub>7</sub> に送出される（図 1 5 のステップ [ 1 ] ）。フロー制御セル F C を受信すると、インタフェース部 2 3 0<sub>7</sub> はバックプレッシャ・バス 7 1 を介して、フロー制御セル F C と同一情報を具備するバックプレッシャ信号 B P 5 を自装置のほか、他の全てのインタフェース部 2 3 0<sub>0</sub> ~ 2 3 0<sub>6</sub> に出力する（ステップ [ 2 ] ）。インタフェース部 2 3 0<sub>7</sub> は、受取ったフロー制御セル F C をインタフェース部 2 2 0<sub>7</sub> に送出する（ステップ [ 3 ] ）。ステップ [ 2 ] でバックプレッシャ信号 B P 5 を受取った各インタフェース部 2 3 0<sub>0</sub> ~ 2 3 0<sub>6</sub> は、対応する F C パケットをインタフェース部 2 2 0<sub>0</sub> ~ 2 2 0<sub>6</sub> に

送出する（ステップ〔4〕）。

#### 固定長セル化に伴うフラグメント効率の悪化対策

可変長パケットを固定長セルに変換（フラグメント化）する際、固定長セルのペイロード部を構成するバイト数に満たないデータはパディングされる。このパディングによって実効スループットが低下し、帯域の無駄使いが発生する。この帯域悪化を防ぐためには、メインスイッチ部232を可変長セルも扱える構成とすることが好ましい。メインスイッチ部232の同一出力ポートを目指すセルが蓄積される出力バッファのセル滞留量に応じて、複数個のセルを結合した形でメインスイッチ部232に送出する（マルチアクセル方式：Multi Adjoining Combined Cell）。

【0067】

今、各入力ポートの物理帯域を $B_{wp}$ とし、セル転送時間を $m$ 、セルとセルの間隔を $n$ とすると、実効帯域 $B_{wa}$ は $B_{wa} = B_{wp} \times m / (m + n)$ となる。マルチアクセルとして $n = 3$ としたトリプルアクセルの場合、 $B_{wa} = B_{wp} \times 3m / (3m + n)$ となり、効率の向上が可能となる。

#### 第1の実施の形態の効果

以下、本発明の第1の実施の形態の主な効果について説明する。

【0068】

第1に、可変長セルを変換して得られた固定長セルをスイッチングするので、各種インタフェースの速度差を効果的に吸収でき、スイッチングによるゆらぎを低減することができる。よって、ATMやEthernetなどの各種インタフェースに対するQoS制御及びバックプレッシャ制御を効果的かつ効率的に行なうことができる。

【0069】

第2に、L3処理部22のインタフェース部220<sub>0</sub>～220<sub>7</sub>の出力側からスイッチ部23のインタフェース部230<sub>0</sub>～230<sub>7</sub>の入力側へのバックプレッシャ信号BP2を、メインスイッチ部232をバイパスするようにしたので、バックプレッシャレイテンシーを低減し、バッファを有効に利用することができる。メインスイッチ部の入出力ポートの帯域を効率的に使用することができる。

## 【 0 0 7 0 】

第 3 に、スイッチ部 2 3 のインタフェース部 2 3 0<sub>0</sub> ~ 2 3 0<sub>7</sub> が輻輳又は輻輳が予期される状態になった時には、インタフェース部 2 2 0<sub>0</sub> ~ 2 2 0<sub>7</sub> の出力側の共通バッファ 3 2 でセルを廃棄するのではなく、入力側の共通バッファ 3 3 でセルを廃棄することとしたため、輻輳発生時のサービス低下を局所化することができる。

## 【 0 0 7 1 】

第 4 に、フロー制御セルを用いたフローレベル、及びバックプレッシャ・バス 7 1 を利用したリンクレベルの 2 段階のバックプレッシャ制御を設けたので、バックプレッシャ制御を効率的かつ効果的に行なうことができる。

第 2 の実施の形態

本発明の第 2 の実施の形態は、前述した従来技術の問題点、つまり異なるネットワークを効率よく収容して効率的な中継処理を行えない問題点を解決するもので、特に先に示した従来技術の第 2 の問題点を解決するものである。

## 【 0 0 7 2 】

図 1 6 は、前述した第 2 の問題点を示す図、図 1 7 はこの問題点を解決する本発明の第 2 の実施の形態を示す図である。

## 【 0 0 7 3 】

図 1 6 ( A ) は、従来の L 3 処理部と回線終端部を示す。L 3 処理部は、スイッチインタフェース部 3 2 4、ローカルスイッチ部 3 2 5 及び回線インタフェース部 3 2 6 を有する。回線終端部は、L 3 インタフェース部 3 1 3 及び物理層処理部 3 1 4 を有する。L 3 処理部は、出力回線単位でバッファを具備しており、出力回線単位にバックプレッシャ制御を行なう。出力回線単位のバックプレッシャ制御なので、前述したように、ルータ全体のバッファを効率的に利用することができない。

## 【 0 0 7 4 】

図 1 6 ( B ) は別の従来例を示す。L 3 処理部は、スイッチインタフェース部 4 2 4、ローカルスイッチ部 4 2 5 及び回線インタフェース部を有する。ローカルスイッチ部 4 2 5 は共通バッファで構成されている。共通バッファ 4 2 5 は、

出力回線（ポート）をアグリゲートする。前述したように、ある出力回線のバックプレッシャ制御の要求を受けると、出力回線が輻輳していないデータまでバックプレッシャの影響を受けてしまい、データを出力できないブロッキング状態が発生してしまう。

#### 【 0 0 7 5 】

図 1 7（A）及び（B）は、上記問題点を解決する構成である。図中、前述した構成要素と同一のものには同一の参照番号を付してある。図 1 7（A）は物理的なバックプレッシャ信号の流れを示し、図 1 7（B）は論理的なバックプレッシャ信号の流れを示す。ローカルスイッチ部 2 2 5 は、大容量のバッファ 2 2 7 を有している。バッファ 2 2 7 は、出力回線単位に内部バッファを具備している。ローカルスイッチ部 2 2 5 は、スイッチインタフェース部 2 2 4 からのセルを出力回線単位に振り分けて、対応する内部バッファにセルを格納する。図 1 7（B）に示すように、出力回線毎のバックプレッシャ信号は、論理的に内部バッファに 1 対 1 に対応している。ローカルスイッチ部 2 2 5 は、バックプレッシャ信号を終端し、出力回線毎にセルを廃棄する。従って、図 1 6（B）に示すようなブロッキングの発生を回避することができる。

### 第 3 の実施の形態

本発明の第 3 の実施の形態は、前述した従来技術の問題点、つまり異なるネットワークを効率よく収容して効率的な中継処理を行えない問題点を解決するもので、特に先に示した従来技術の第 3 の問題点を解決するものである。

#### 【 0 0 7 6 】

図 1 8 は、本発明の第 3 の実施の形態を示す図である。図中、前述した構成要件と同一のものには同一の参照番号を付してある。二重化構成においては、運用系のバックプレッシャか非運用系のバックプレッシャを判別することが好ましい。非運用系からのバックプレッシャ制御の要求を受けた運用系は、セル送信を停止することをしないで、この要求を破棄する。運用系か非運用系かの判断は、例えば装置状態を示す信号を参照する。つまり、バックプレッシャ信号の中に運用系／非運用系を示すフラグを設けておけば良い。

#### 【 0 0 7 7 】

図 1 8 に示すように、非運用系からのバックプレッシャ信号は運用系で破棄されている。これに対し、運用系から非運用系へのバックプレッシャ信号は非運用系で破棄されず、各ポイントで受け付けられる。

#### 第 4 の実施の形態

本発明の第 4 の実施の形態は、前述した従来技術の問題点、つまり異なるネットワークを効率よく収容して効率的な中継処理を行えない問題点を解決するもので、特に先に示した従来技術の第 4 の問題点を解決するものである。

##### 【 0 0 7 8 】

図 1 9 は、本発明の第 4 の形態を示す図である。図 1 9 において、前述した構成要素と同一のものには同一の参照番号を付してある。回線終端部 2 1 は、バッファ容量監視部 2 1 5 を有する。バッファ容量監視部 2 1 5 は、物理層処理部 2 1 4 内部の回線対応（ポート対応）の内部バッファの容量を監視しており、監視結果を直接ローカルスイッチ部 2 2 7 に知らせる。また、回線インタフェース部 2 2 6 及び L 3 インタフェース部 2 1 3 をバッファレスの構成とする。これにより、ローカルスイッチ部 2 2 5 と物理層処理部 2 1 4 との間でセルをバッファリングしないので、バックプレッシャ制御が要求された場合でもバッファリングされることがない。バックプレッシャ制御の解除後もセルはバッファリングされないで、バーストで出力回線へデータが送出されることがない。

##### 【 0 0 7 9 】

このようにして、セルのゆらぎや遅延を最小限に抑えることができる。

##### 【 0 0 8 0 】

以上、本発明の実施の形態を説明した。本発明は、上述した第 1 ないし第 4 の実施の形態に限定されるものではなく、他の様々の実施の形態を含むものである。

##### 【 0 0 8 1 】

以下、本発明の特徴の一部を列挙する。

##### 【 0 0 8 2 】

（付記 1） 異なるインタフェース間のスイッチング機能を有する通信装置において、

固定長のデータをスイッチングするメインスイッチ部と、該メインスイッチ部の入力及び出力にそれぞれ設けられた第 1 及び第 2 のバッファを具備するインタフェース部とを設けたスイッチ部を有することを特徴とする通信装置。

【 0 0 8 3 】

(付記 2) 前記通信装置は更に、前記スイッチ部に接続され、所定のプロトコルに従った処理を行う処理部を有し、

該処理部は前記第 1 及び第 2 のバッファに接続する第 3 及び第 4 のバッファを有し、

前記第 1 のバッファが所定の状態になった時に、前記第 3 のバッファに対しバックプレッシャ制御を行うことを特徴とする付記 1 記載の通信装置。

【 0 0 8 4 】

(付記 3) 前記通信装置は更に、前記スイッチ部に接続され、所定のプロトコルに従った処理を行う処理部を有し、

該処理部は前記第 1 及び第 2 のバッファに接続する第 3 及び第 4 のバッファを有し、

前記第 4 のバッファが所定の状態になった時に、前記第 1 のバッファに対しバックプレッシャ制御を行なうことを特徴とする付記 1 記載の通信装置。

【 0 0 8 5 】

(付記 4) 前記第 2 のバッファが所定の状態になった時に、前記第 1 のバッファに対しバックプレッシャ制御を行うことを特徴とする付記 1 記載の通信装置。

【 0 0 8 6 】

(付記 5) 前記通信装置は更に、前記スイッチ部に接続され、所定のプロトコルに従った処理を行う処理部を有し、

該処理部は前記第 1 及び第 2 のバッファに接続する第 3 及び第 4 のバッファを有し、

前記処理部に接続された装置からのバックプレッシャ制御の要求を受けると、前記第 4 のバッファに対しバックプレッシャ制御を行うことを特徴とする付記 1 記載の通信装置。

【 0 0 8 7 】

(付記 6) 前記バックプレッシャ制御の要求は、所定のフロー制御セルで形成されることを特徴とする付記 2 ないし 4 のいずれか一項記載の通信装置。

【 0 0 8 8 】

(付記 7) 前記バックプレッシャ制御は、所定の Q o S クラス単位に行なわれることを特徴とする付記 2 ないし 5 のいずれか一項記載の通信装置。

【 0 0 8 9 】

(付記 8) 前記バックプレッシャ制御は、回線単位に行なわれることを特徴とする付記 2 ないし 5 のいずれか一項記載の通信装置。

【 0 0 9 0 】

(付記 9) 前記所定の状態は、所定の Q o S クラス単位に判断することを特徴とする付記 2 ないし 4 のいずれか一項記載の通信装置。

【 0 0 9 1 】

(付記 1 0) 前記所定の状態は、回線単位に判断することを特徴とする付記 2 ないし 4 のいずれか一項記載の通信装置。

【 0 0 9 2 】

(付記 1 1) 前記処理部は、前記スイッチ部から受取ったデータを回線対応の内部バッファに供給するローカルスイッチを有することを特徴とする付記 2 記載の通信装置。

【 0 0 9 3 】

(付記 1 2) 前記処理部は、前記スイッチ部から受取ったデータを一旦格納するバッファを具備したローカルスイッチを有し、該ローカルスイッチは該バッファからデータを読み出して回線対応の内部バッファに供給するローカルスイッチを有することを特徴とする付記 2 記載の通信装置。

【 0 0 9 4 】

(付記 1 3) 前記通信装置は前記処理部と回線との間に設けられた終端部を有し、

該終端装置は回線毎に設けられたバッファと、該バッファの容量を監視して前記スイッチ部から受取ったデータを一旦格納するバッファを制御する監視部とを

有することを特徴とする付記 1 2 記載の通信装置。

【 0 0 9 5 】

(付記 1 4) 前記スイッチ部は二重化構成であり、非運用系からのバックプレッシャ制御の要求を受けた運用系は、バックプレッシャ制御の要求を破棄することを特徴とする付記 1 ないし 1 2 のいずれか一項記載の通信装置。

【 0 0 9 6 】

(付記 1 5) 異なるインタフェース間のスイッチング機能を有する通信装置において、

固定長のデータをスイッチングするメインスイッチ部と、該メインスイッチ部の入力及び出力にそれぞれ設けられた第 1 及び第 2 のバッファを回線毎に複数個具備するインタフェース部とを有するスイッチ部を複数有することを特徴とする通信装置。

【 0 0 9 7 】

(付記 1 6) 前記通信装置は更に、前記スイッチ部に接続され、所定のプロトコルに従った処理を行う複数の処理部と、該複数の処理部に対し所定のバックプレッシャ信号を伝送するバスとを有することを特徴とする付記 1 5 記載の通信装置。

【 0 0 9 8 】

(付記 1 7) 前記複数のスイッチ部の各々は、前記バスを介してバックプレッシャ信号を受け取り、前記第 1 のバッファの読み出しを制御することを特徴とする付記 1 6 記載の通信装置。

【 0 0 9 9 】

(付記 1 8) 異なるインタフェース間のスイッチングを行なう通信制御方法において、

前記異なるインタフェースで扱われるデータに関連する固定長のデータを一旦バッファリングしてからスイッチングする段階と、

スイッチングされたデータを一旦バッファリングしてから、回線に向けて送出する段階と

を有する通信制御方法。



【 0 1 0 0 】

(付記 1 9) スイッチングする前のバッファリングが所定の状態になった時は、バックプレッシャ制御の要求をスイッチングをバイパスして、他の装置に送出することを特徴とする付記 1 8 記載の通信制御方法。

【 0 1 0 1 】

【発明の効果】

以上説明したように、本発明によれば、異なるネットワークを効率よく収容して効率的に中継処理を行える通信装置を提供することができる。

【図面の簡単な説明】

【図 1】

ネットワーク構成の一例を示す図である。

【図 2】

本発明の第 1 の実施の形態によるルータの構成を示すブロック図である。

【図 3】

図 2 の各部の内部構成例を示すブロック図である。

【図 4】

レイヤ 2 の終端処理とレイヤ 3 のルーティング処理を示す図である。

【図 5】

L 3 処理部 2 2 及びスイッチ部 2 3 の基本構成を示すブロック図である。

【図 6】

第 1 のスケジューラの一構成例を示す図である。

【図 7】

第 2 のスケジューラの一構成例を示す図である。

【図 8】

インタフェース部の共通バッファ及び出力バッファの一構成例を示す図である。

【図 9】

第 3 のスケジューラの一構成例を示す図である。

【図 1 0】

バックプレッシャ信号によるバックプレッシャ制御を説明するための図である

【図 1 1】

インタフェース部の入力バッファと共通バッファを模式的に図示してある。

【図 1 2】

バックプレッシャ信号によるバックプレッシャ制御を説明するための図である

【図 1 3】

バックプレッシャ信号の生成の一例を示す図である。

【図 1 4】

別のバックプレッシャ信号の生成の一例を示す。

【図 1 5】

L 3 処理部の回線対応処理部内部のインタフェース部に設けられた出力バッファ内のあるクラスに対応するキュー内のセル数又は流量が所定のしきい値を超えた場合を示す。

【図 1 6】

第 2 の問題点を示す図である。

【図 1 7】

本発明の第 2 の実施の形態を示す図である。

【図 1 8】

本発明の第 3 の実施の形態を示す図である。

【図 1 9】

本発明の第 4 の実施の形態を示す図である。

【符号の説明】

2 2      L 3 処理部

2 2 <sub>0</sub> ～ 2 2 <sub>8</sub>      回線対応処理部

2 3      スイッチ部

3 1 ～ 3 4      共通バッファ

3 5、3 8      入力バッファ

3 6、3 7 出力バッファ

2 2 0<sub>0</sub> ~ 2 2 0<sub>8</sub> インタフェース部

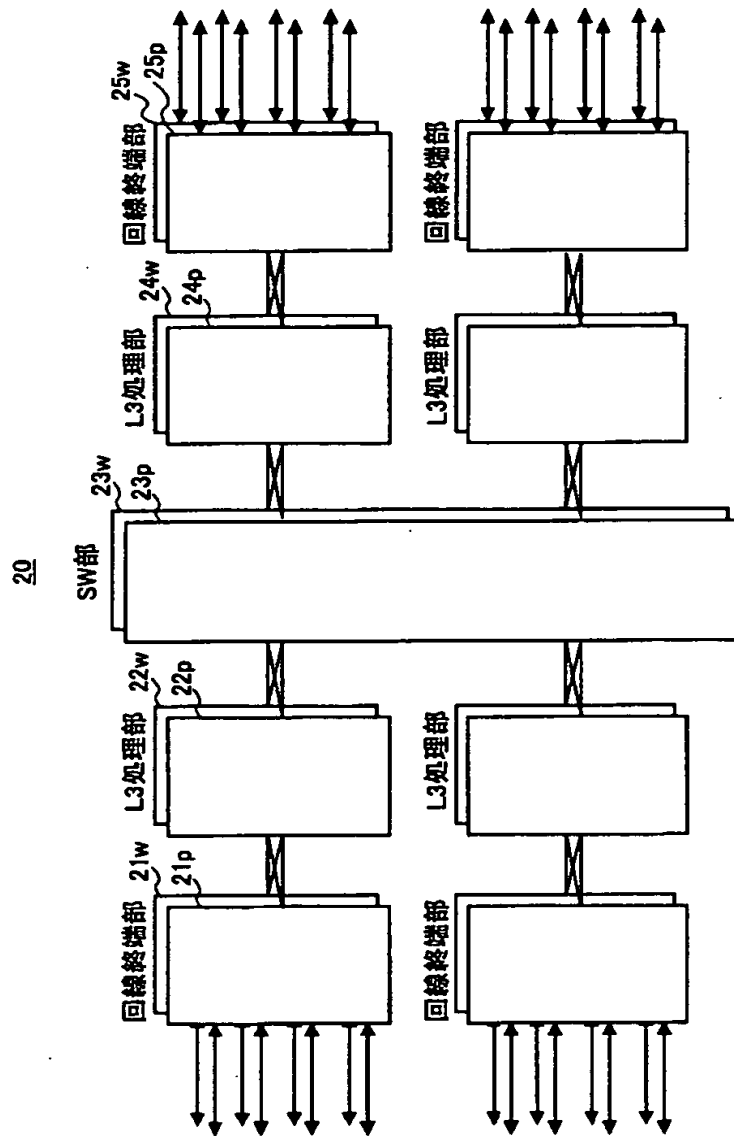
2 3 0<sub>0</sub> ~ 2 3 0<sub>7</sub> インタフェース部

2 3 2 メインスイッチ部



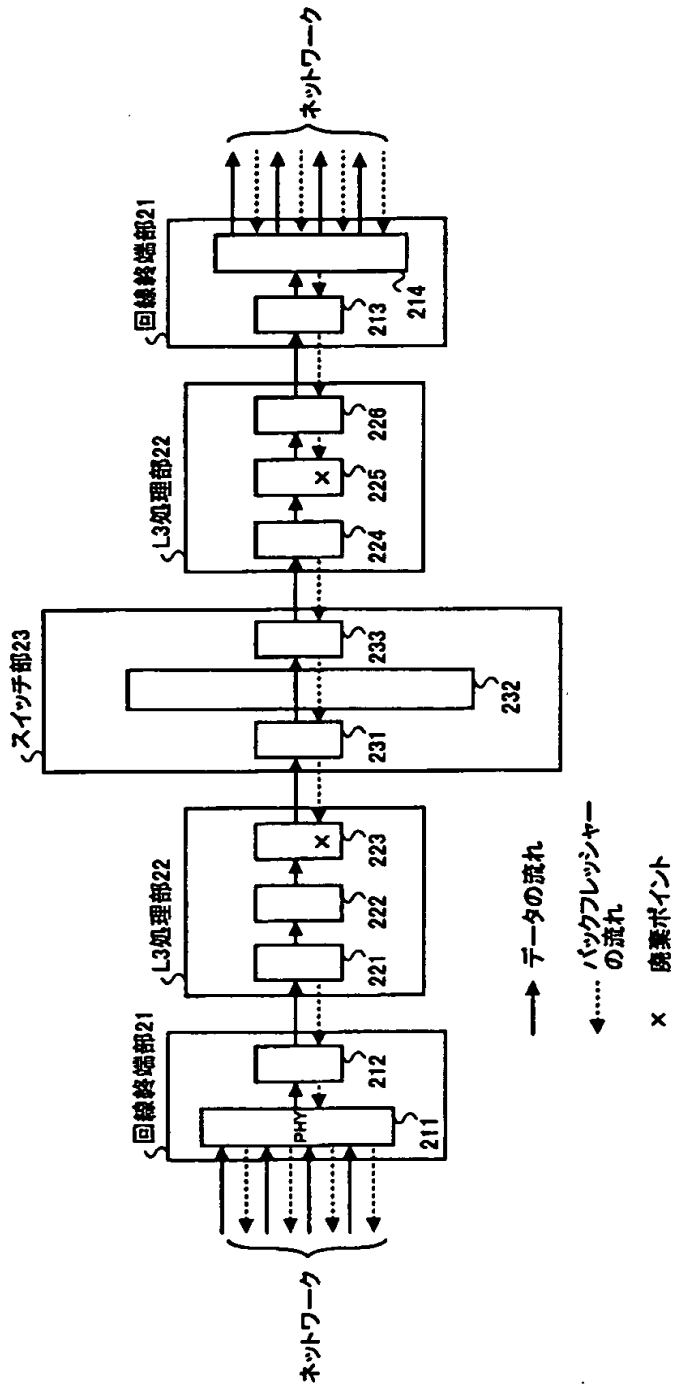
【図 2】

本発明の第1の実施の形態によるルータの構成を示すブロック図



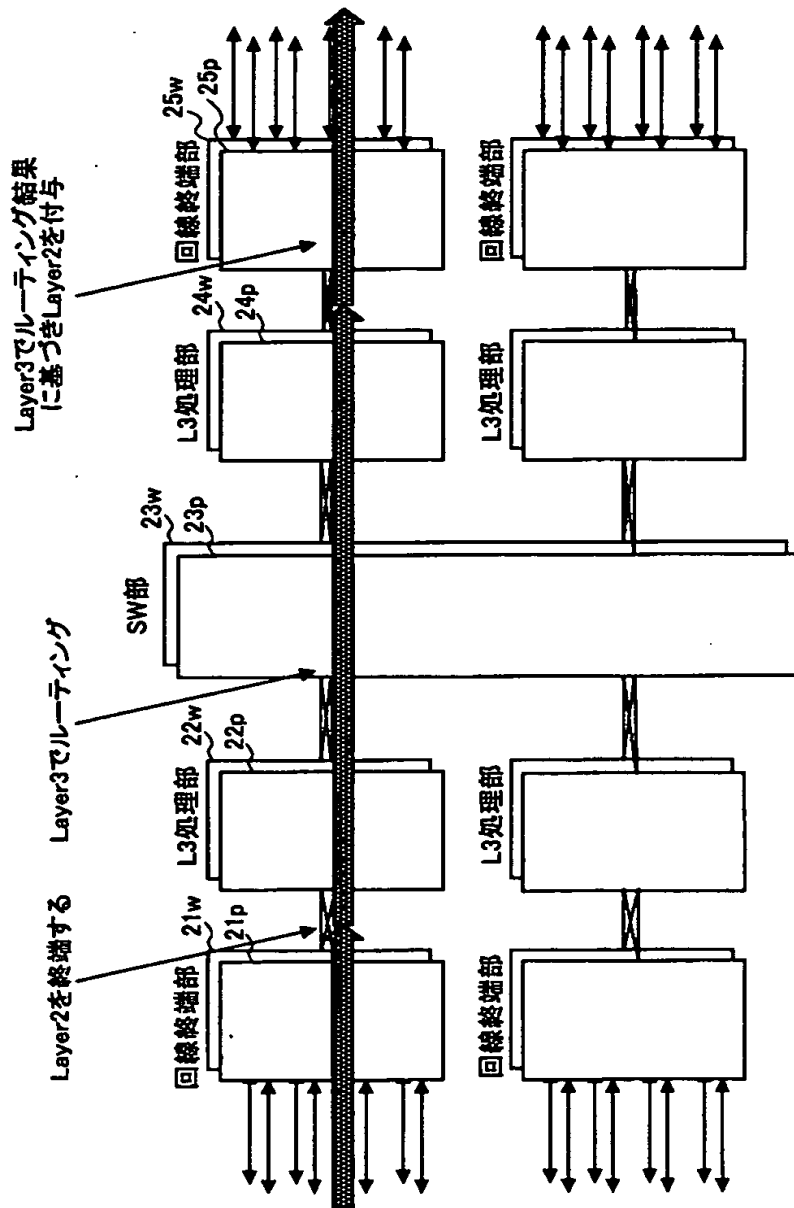
【図 3】

図2の各部の内部構成例を示すブロック図



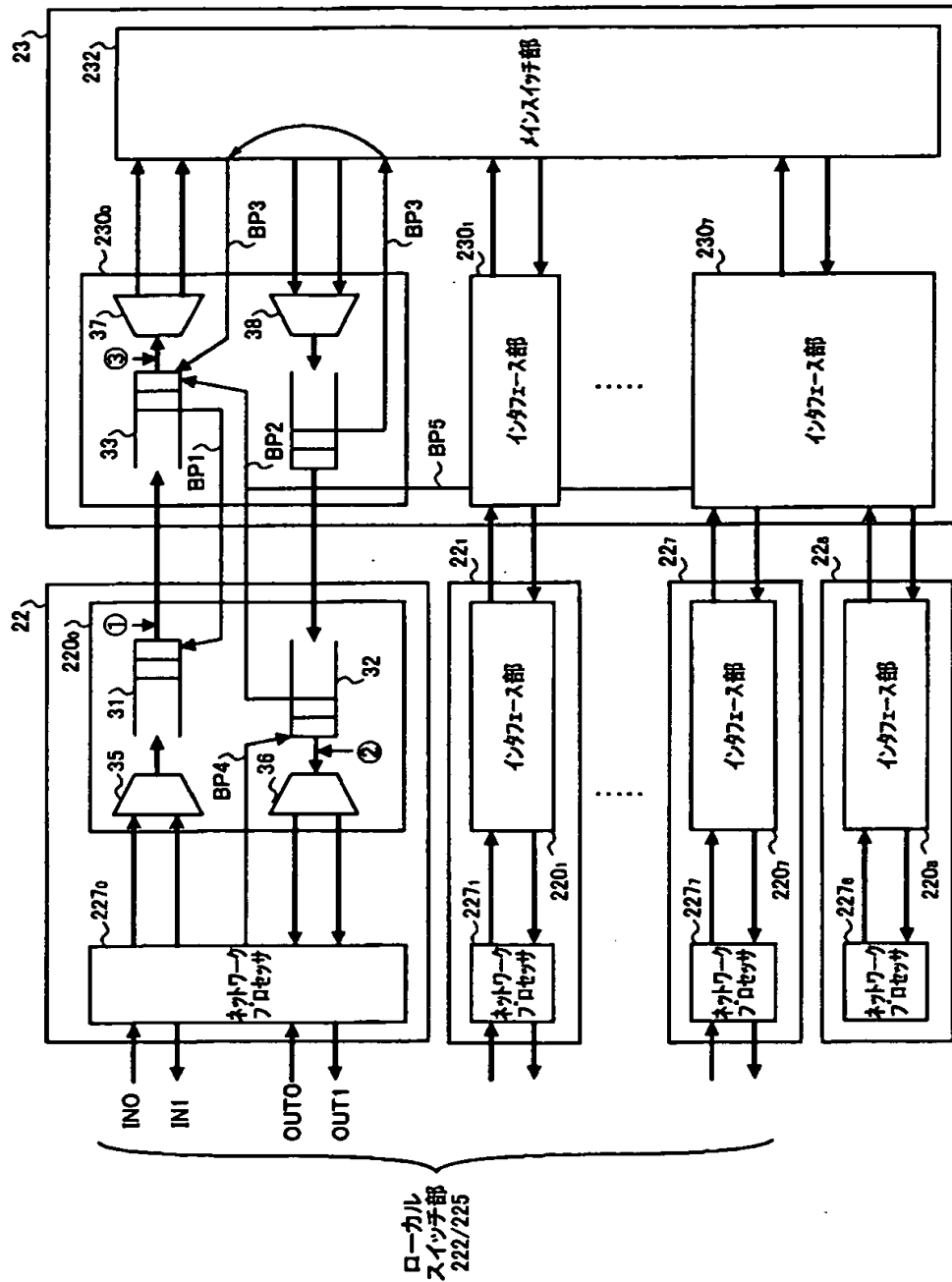
【図 4】

レイヤ2の終端処理とレイヤ3のルーティング処理を示す図



【図5】

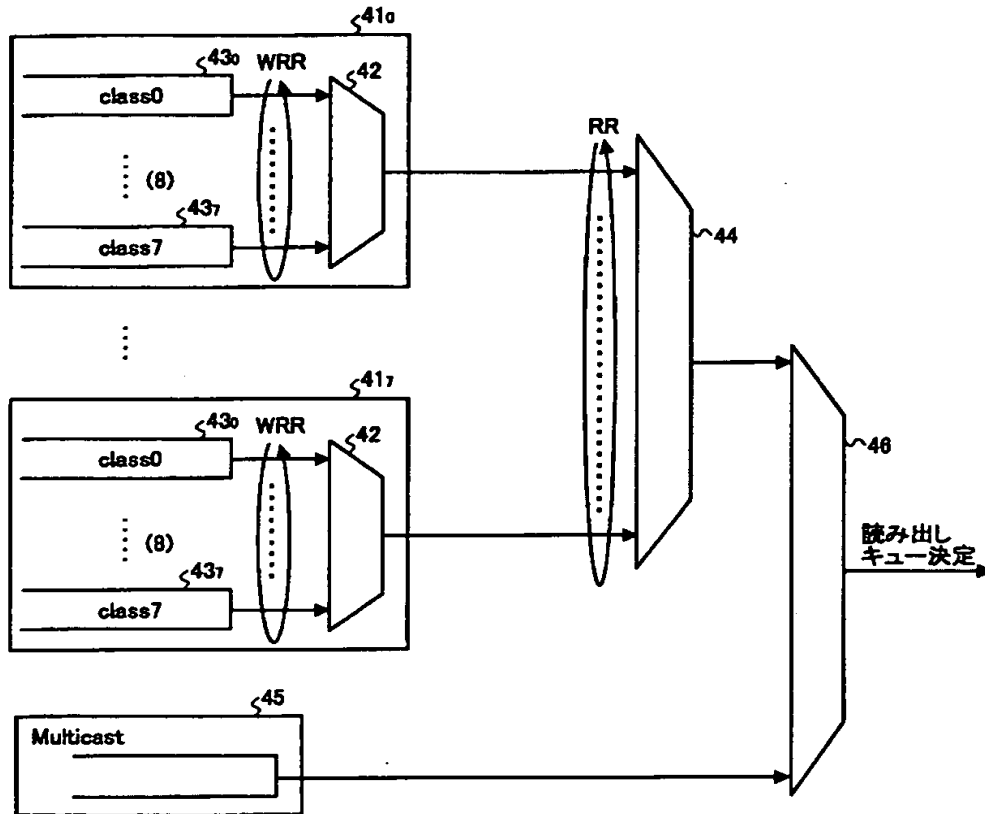
L3処理部22及びスイッチ部23の基本構成を示すブロック図





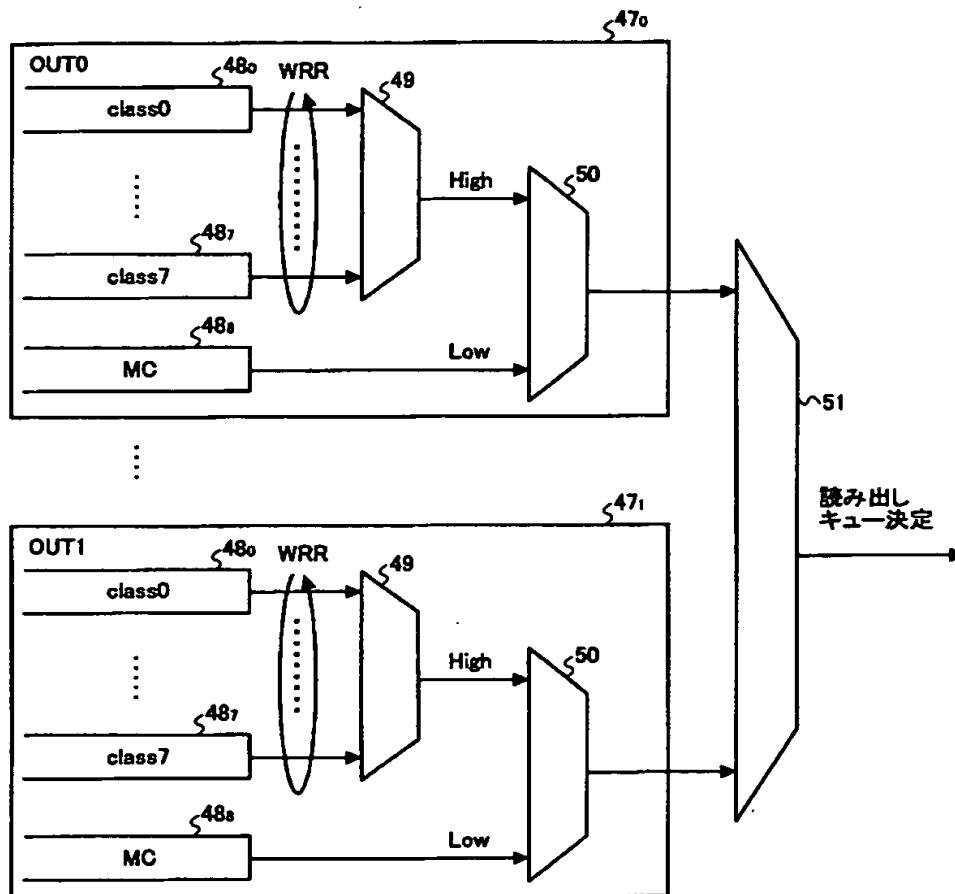
【図 6】

第1のスケジューラの一構成例を示す図



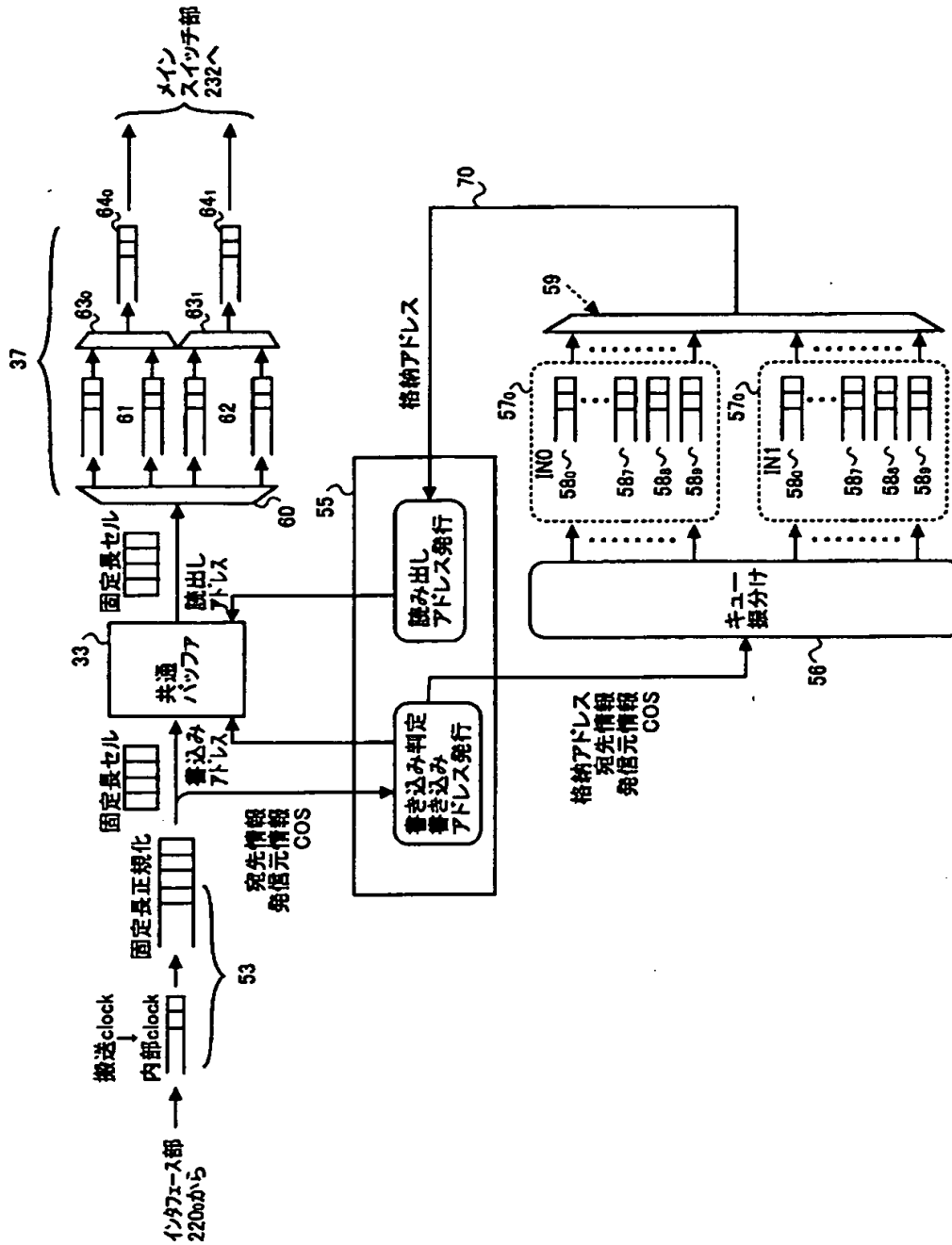
【図 7】

第2のスケジューラの一構成例を示す図



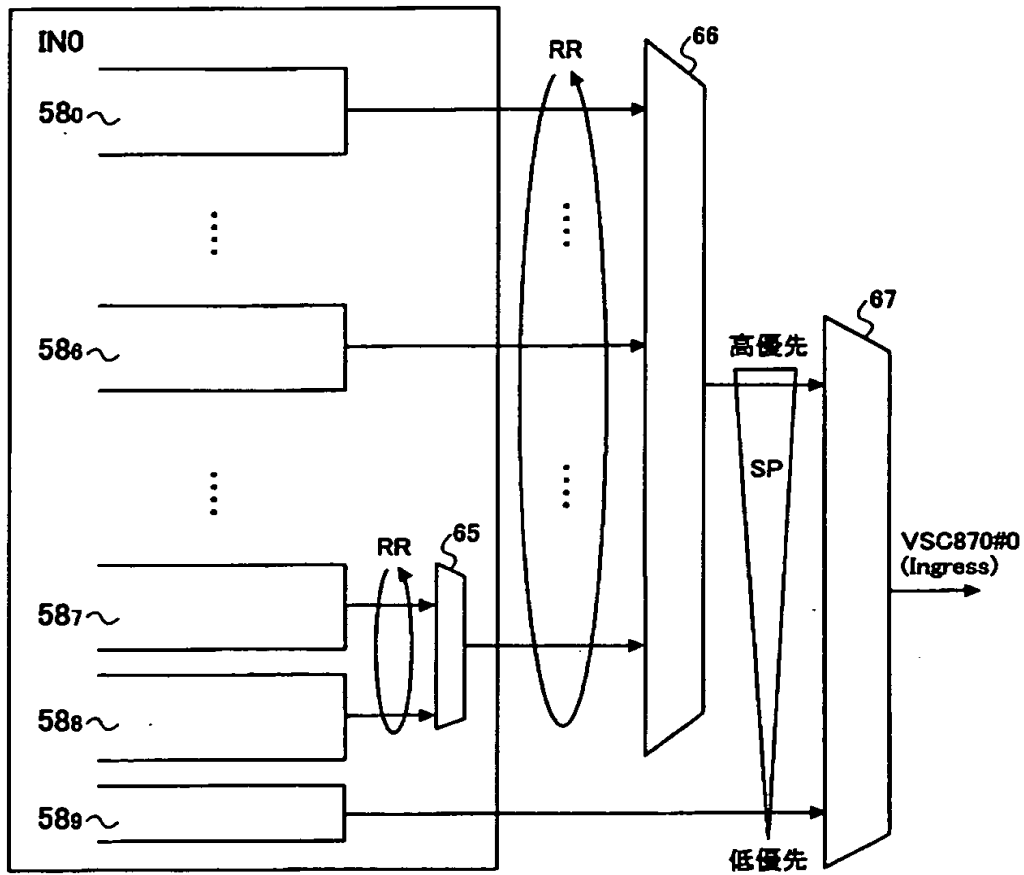
【図 8】

インタフェース部の共通バッファ及び  
出力バッファの一構成例を示す図



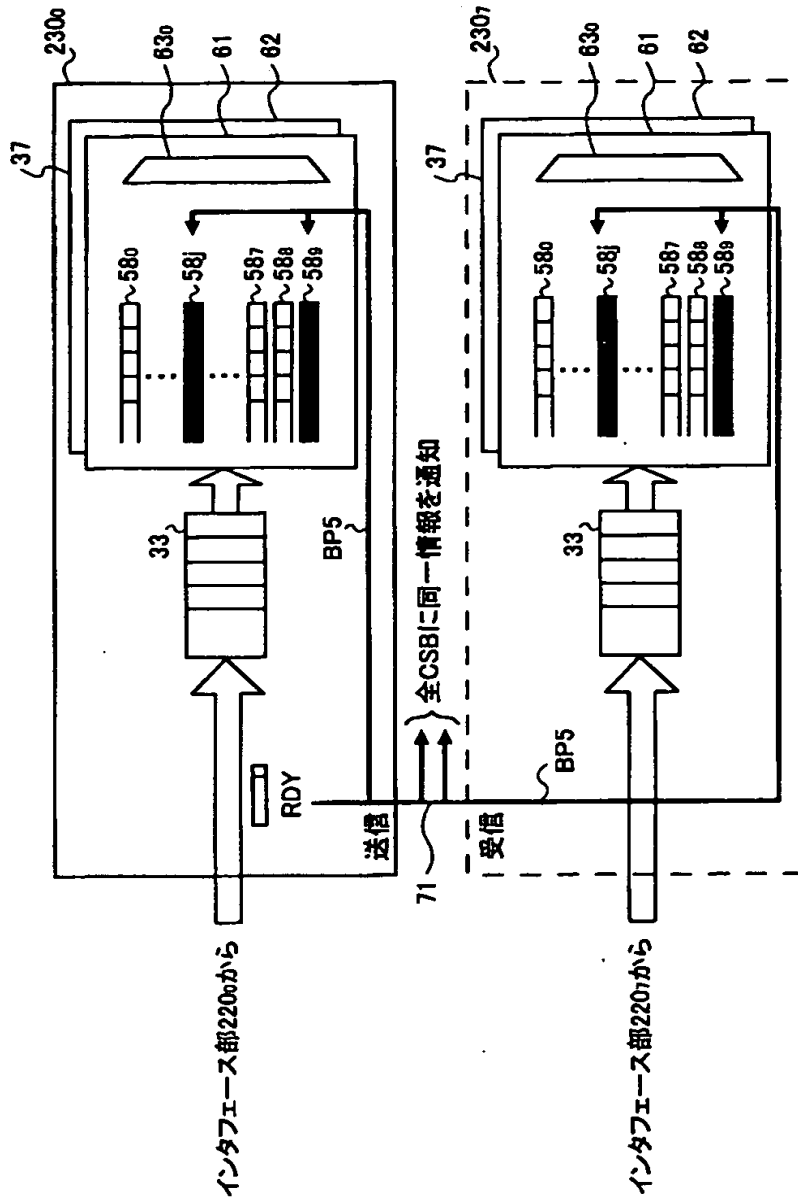
【図 9】

第3のスケジューラの一構成例を示す図



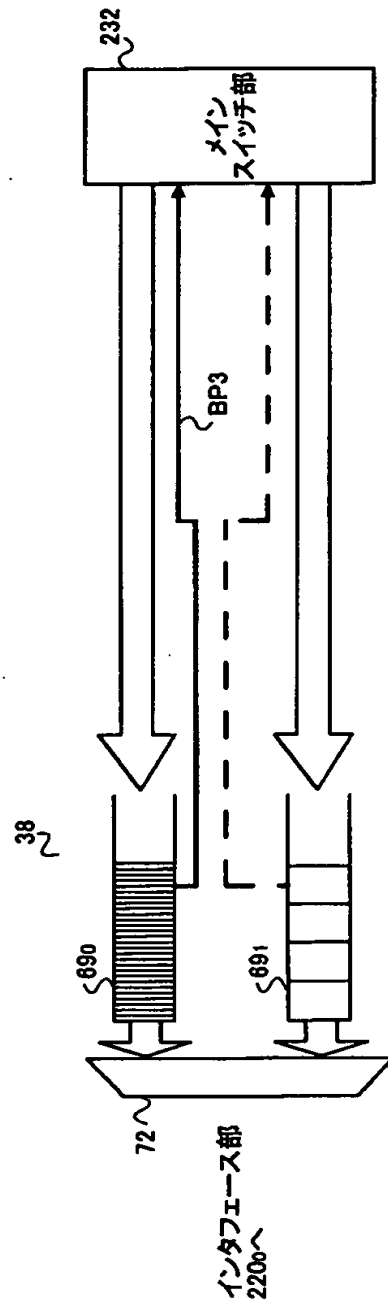
【図10】

バックプレッシャ信号による  
バックプレッシャ制御を説明するための図



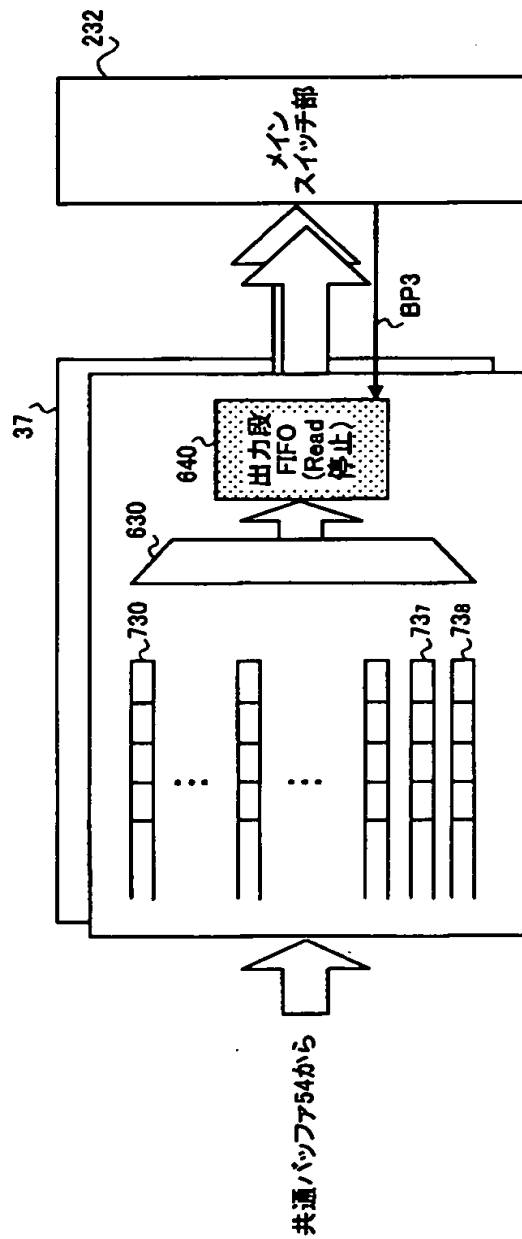
【図 11】

インタフェース部の入力バッファと共通バッファを模式的に示してある図



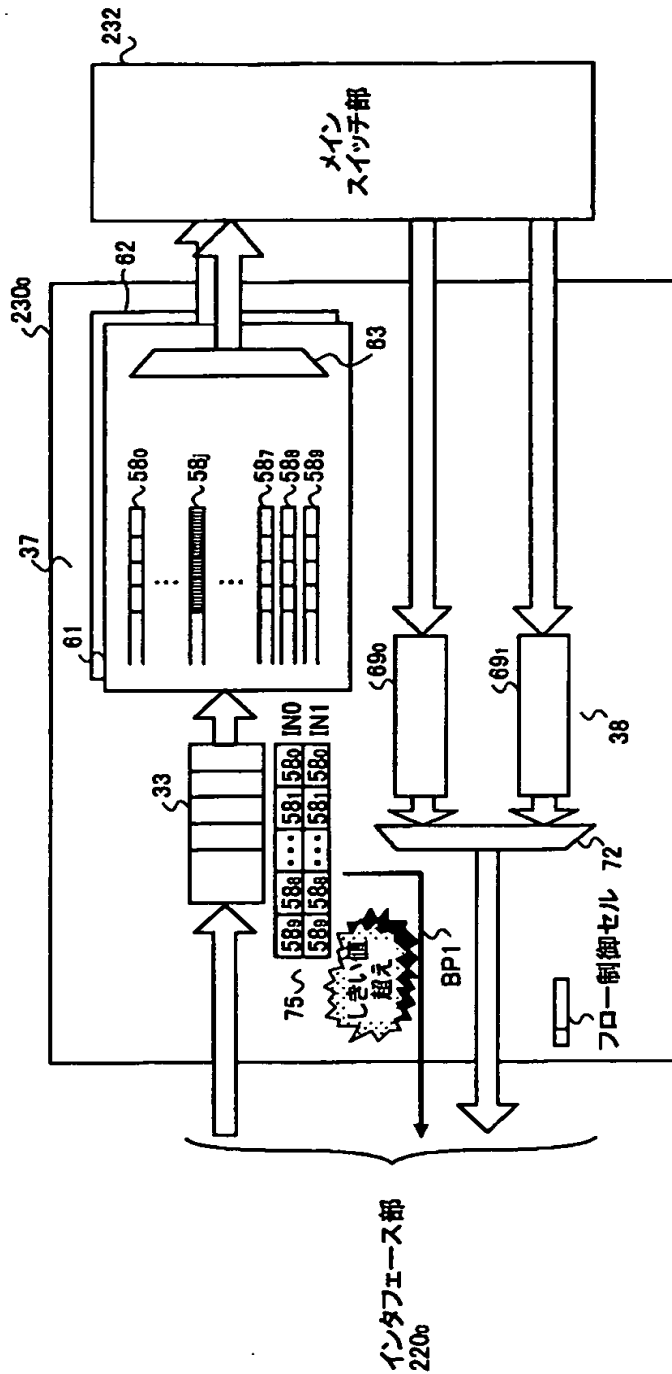
【図 12】

バックプレッシャ信号による  
バックプレッシャ制御を説明するための図



【図 13】

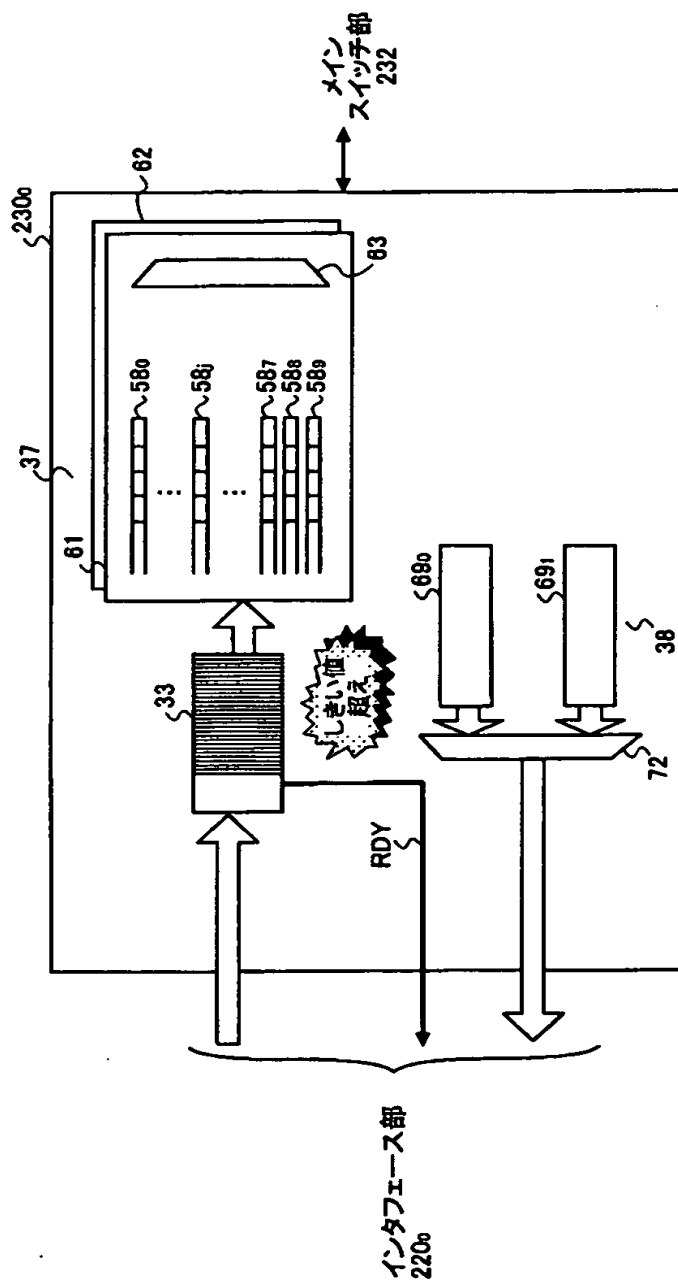
バックプレッシャ信号の生成の一例を示す図





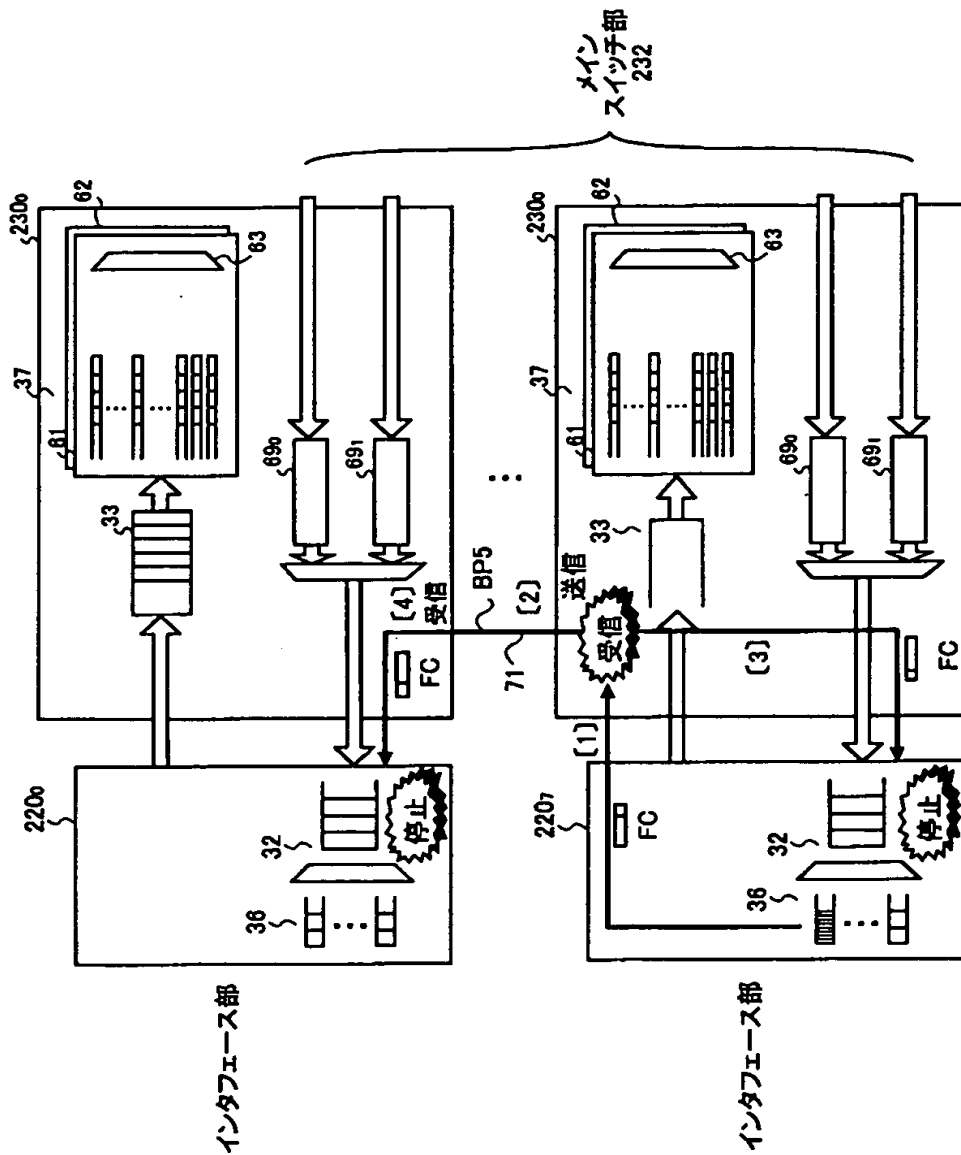
【図 14】

別のバックプレッシャ信号の生成の一例を示す図



【図15】

L3処理部の回線対応処理内部のインタフェース部に設けられた  
出力バッファ内のあるクラスに対応するキュー内のセル数  
又は流量が所定のしきい値を越え場合を示す図

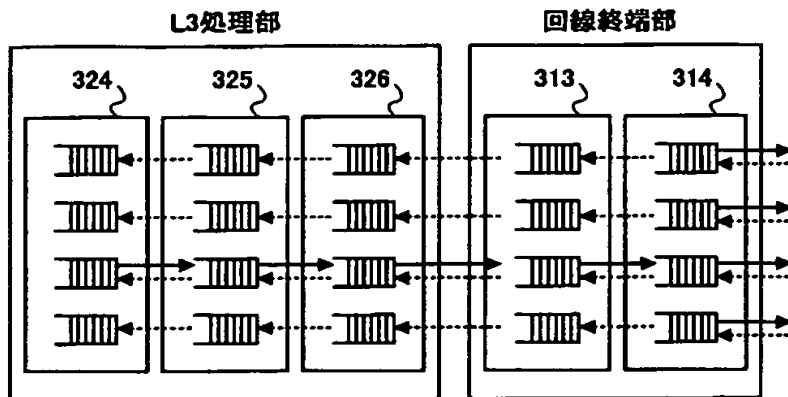


【図 16】

第2の問題点を示す図

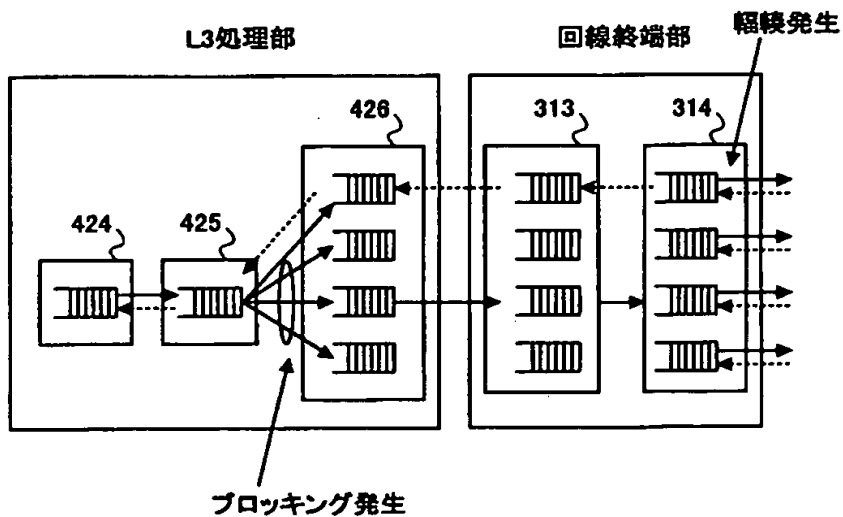
(A)

複数のバッファを具備して構成する場合



(B)

ブロッキングが発生してしまう場合



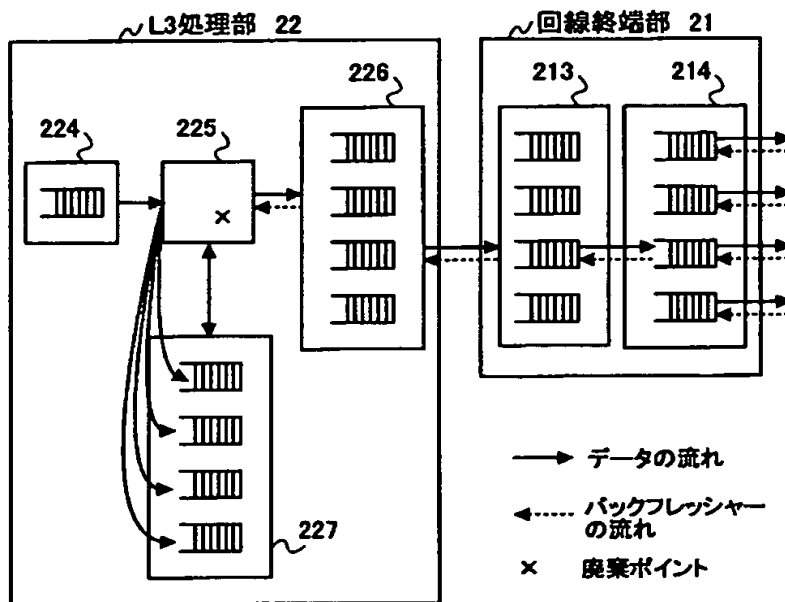
→ データの流れ  
 ←..... バックプレッシャー  
 の流れ  
 × 廃棄ポイント

【図 17】

本発明の第2の実施の形態を示す図

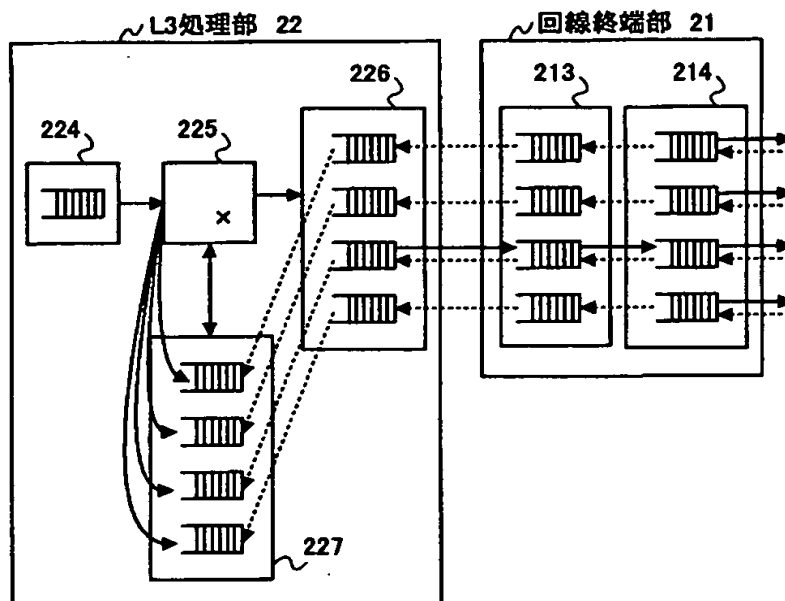
(A)

物理的なバックプレッシャーの流れ



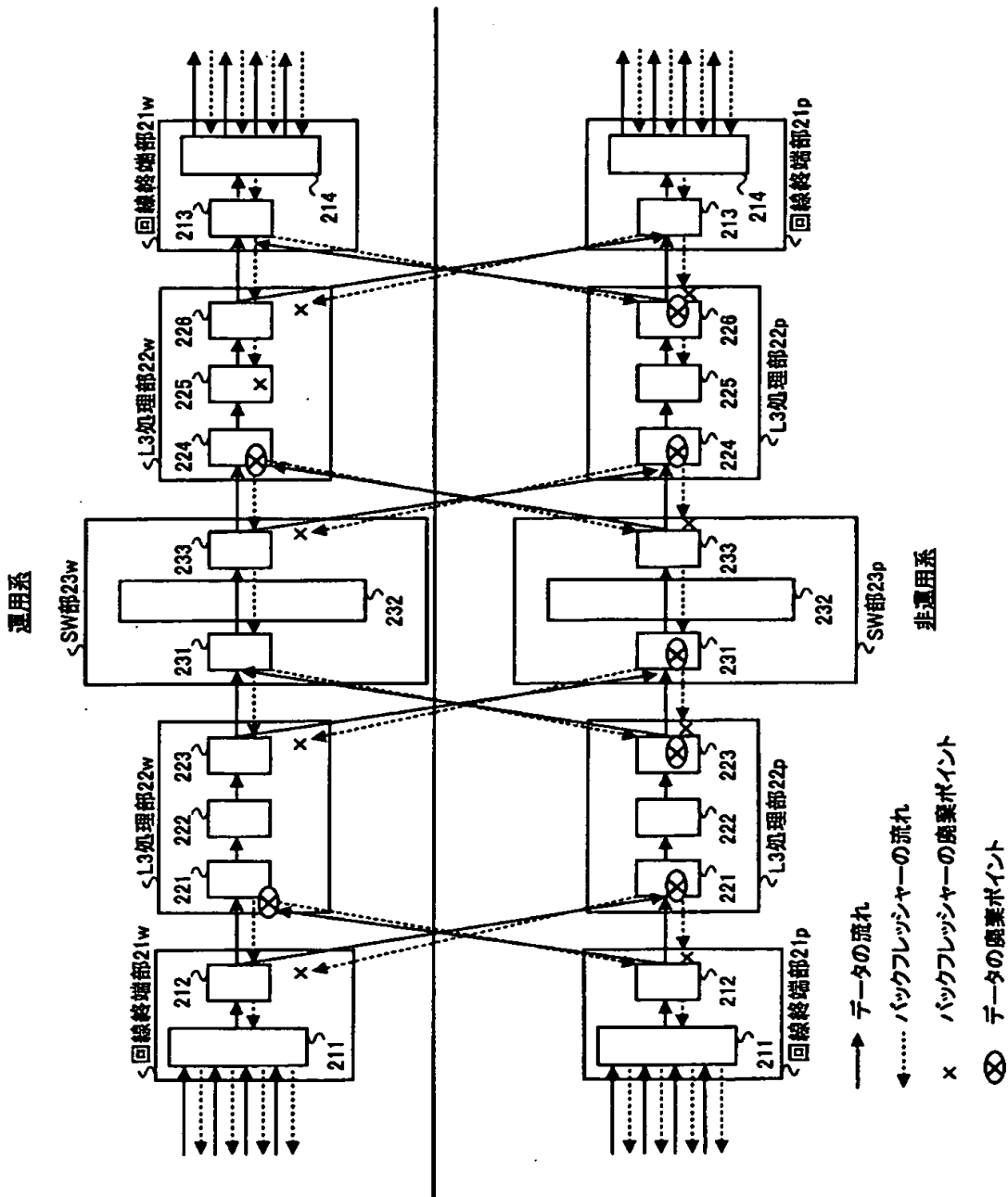
(B)

理論的なバックプレッシャーの流れ



【図18】

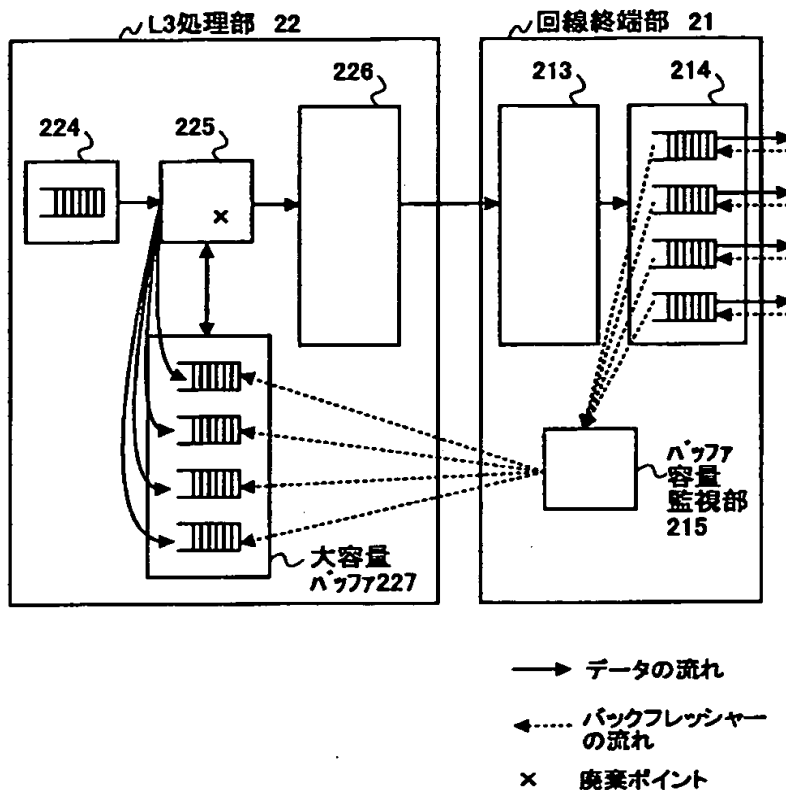
本発明の第3の実施の形態を示す図



【図 1 9】

本発明の第4の実施の形態を示す図

ATM回線時の理論的なバックプレッシャーの流れ



【書類名】 要約書

【要約】

【課題】 異なるネットワークを効率よく収容して効率的に中継処理を行える通信装置を提供する

【解決手段】 固定長のデータをスイッチングするメインスイッチ部（232）と、該メインスイッチ部の入力及び出力にそれぞれ設けられた第1のバッファ（33、37）及び第2のバッファ（34、38）を具備するインタフェース部（230<sub>0</sub>）とを有するスイッチ部を有する。

【選択図】 図5

特2001-196778

出 願 人 履 歴 情 報

識別番号 [000005223]

1. 変更年月日 1996年 3月26日

[変更理由] 住所変更

住 所 神奈川県川崎市中原区上小田中4丁目1番1号

氏 名 富士通株式会社